

From: Leffers, Gerald
Sent: Tuesday, January 25, 2005 4:06 PM
To: STIC-Biotech/ChemLib
Subject: FW:10/776,213 sequence search

Examiner's mailbox is in 2c70 of Remsen. Thank you. Gerry Leffers

Gerald G. Leffers Jr., PhD
Primary Examiner, Art Unit 1636
Remsen Building, Room 02A69
(571) 272-0772

-----Original Message-----

From: Leffers, Gerald
Sent: Tuesday, January 25, 2005 4:02 PM
To: STIC-Biotech/ChemLib
Subject: 09/776,213 sequence search

Please do a search/interference search for SEQ ID NO: 2 of this application (~723 nucleotides). Claims are to sequences comprising as few as 17 consecutive nucleotides of SEQ ID NO: 2. Thank you. Gerry Leffers

Gerald G. Leffers Jr., PhD
Primary Examiner, Art Unit 1636
Remsen Building, Room 02A69
(571) 272-0772

RECEIVED
JAN 25 2005
STIC-Biotech/ChemLib
(STIC)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)


```
Db      241 GCACTAGTACTAATGNGCTGGAACATCTCCATGGCTGTGACTTGTGTCTATCTCAAGT 300
Qy      301 GGTAAACGGCAACCGTGGCTCGGAAACGTTTCTTGTGCAAAATTTAGAAACAGGGGCTACA 360
Db      301 GGTAAACGGCAACCGTGGCTCGGAAACGTTTCTTGTGCAAAATTTAGAAACAGGGGCTACA 360
Qy      361 GTCCTGATATAGAAATTAAGCGCATTTTGTGCTAGCGCGCGCGCGCGCGCGCGCTTCCCA 420
Db      361 GTCCTGATATAGAAATTAAGCGCATTTTGTGCTAGCGCGCGCGCGCGCGCGCGCTTCCCA 420
Qy      421 ATAGGAGGCGCAGTTTATCGCGCGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 480
Db      421 ATAGGAGGCGCAGTTTATCGCGCGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 480
Qy      481 TGTTCGCGCCAGTGTGCTGCAAGGCTGCGCGCGGAACTAGTATAGAGGATGTAAAC 540
Db      481 TGTTCGCGCCAGTGTGCTGCAAGGCTGCGCGCGGAACTAGTATAGAGGATGTAAAC 540
Qy      541 TTTGATGAGAAATTAGCAAGCGGAAAAAACTATGCGCTAGCGGAGTTGTTTCAA 600
Db      541 TTTGATGAGAAATTAGCAAGCGGAAAAAACTATGCGCTAGCGGAGTTGTTTCAA 600
Qy      601 TCATATTAAGGGAATTTGCTCTCATATGTGACAGTTTCTGGACGCTTAACTTT 660
Db      601 TCATATTAAGGGAATTTGCTCTCATATGTGACAGTTTCTGGACGCTTAACTTT 660
Qy      661 TATTGACAGAGACTATCAATCATACAGATTTCTCAAAAAAAGACTAATATTA 720
Db      661 TATTGACAGAGACTATCAATCATACAGATTTCTCAAAAAAAGACTAATATTA 720
Qy      721 CAT 723
Db      721 CAT 723

RESULT 2
AR492049      11427 bp      DNA      linear      PAT 15-MAY-2004
LOCUS      AR492049      Sequence 24 from patent US 6716601.
DEFINITION      AR492049
ACCESSION      AR492049.1 GI:47260518
VERSION
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unclassified.
REFERENCE      1 (bases 1 to 11427)
AUTHORS      Belfield,G.P. and Oakley,C.
TITLE      Compositions and methods utilizing the yeast ZEO1 promoter
JOURNAL      Patent: US 6716601-A 24 06-APR-2004;
FEATURES
     source      location/Qualifiers
                1..11427
                /organism="unknown"
                /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 723; DB 6; Length 11427;
Best Local Similarity 100.0%; Pred. No. 6.1e-207;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTCGATTAGCAGCAGACATCATAGACTGGCTGATAAAATATACATACGAAAA 60
Db      15 CTTTCGATTAGCAGCAGACATCATAGACTGGCTGATAAAATATACATACGAAAA 74
Qy      61 ACCATAAAGACAAAGCATCTACTTGTGAAGAAAGAGACGCTTGTGAAGGGGAT 120
Db      75 ACCATAAAGACAAAGCATCTACTTGTGAAGAAAGAGACGCTTGTGAAGGGGAT 134
Qy      121 GGGGGCTAAGAAATCACTTCTTTTCCCTGCGGCTCGGACCGGGGACCCCTCT 180
Db      135 GGGGGCTAAGAAATCACTTCTTTTCCCTGCGGCTCGGACCGGGGACCCCTCT 194
Qy      181 CTCGCCGACGATTTCTTCTTATATCTCTTATATCTTATTCGATCCGTTGAAGAAC 240
```

```
Db      195 CTCGCCGACGATTTCTTCTTATATCTTCTTATATCTTATTCATCCCGTTGAAGAAC 254
Qy      241 GCACTATGACTTAATGGTGTGACATCTCCAGGCGTGAAGCTTGTGTATCTCAAGT 300
Db      255 GCACTATGACTTAATGGTGTGACATCTCCAGGCGTGAAGCTTGTGTATCTCAAGT 314
Qy      301 GGTAAACGGCAACCGTGGCTCGGAAACGTTTCTTGTGACAAATTTAGAAACAGGGGCTACA 360
Db      315 GGTAAACGGCAACCGTGGCTCGGAAACGTTTCTTGTGACAAATTTAGAAACAGGGGCTACA 374
Qy      361 GTCCTGATATAGAAATTAAGCGCAATTTTGTAGCGCGCGCGCGCGCGCGCTTCCCA 420
Db      375 GTCCTGATATAGAAATTAAGCGCAATTTTGTAGCGCGCGCGCGCGCGCGCTTCCCA 434
Qy      421 ATAGGAGGCGCAGTTTATCGCGGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 480
Db      435 ATAGGAGGCGCAGTTTATCGCGGAGCTCTACTTCTTCTATTTGGGTAAGCCCTTTC 494
Qy      481 TGTTCGCGCCAGTGTGCTGCAAGGCTGCGCGCGGAACTAGTATAGAGGATGTAAAC 540
Db      495 TGTTCGCGCCAGTGTGCTGCAAGGCTGCGCGCGGAACTAGTATAGAGGATGTAAAC 554
Qy      541 TTTGATGAGAAATTAGCAAGCGGAAAAAACTATGCGCTAGCGGAGTTGTTTCAA 600
Db      555 TTTGATGAGAAATTAGCAAGCGGAAAAAACTATGCGCTAGCGGAGTTGTTTCAA 614
Qy      601 TCATATTAAGGGAATTTGCTCTCATATGTGACAGTTTCTGGACGCTTAACTTT 660
Db      615 TCATATTAAGGGAATTTGCTCTCATATGTGACAGTTTCTGGACGCTTAACTTT 674
Qy      661 TATTGACAGAGACTATCAATCATACAGATTTCTCAAAAAAAGACTAATATTA 720
Db      675 TATTGACAGAGACTATCAATCATACAGATTTCTCAAAAAAAGACTAATATTA 734
Qy      721 CAT 723
Db      735 CAT 737

RESULT 3
AR492045      13073 bp      DNA      linear      PAT 15-MAY-2004
LOCUS      AR492045      Sequence 20 from patent US 6716601.
DEFINITION      AR492045
ACCESSION      AR492045.1 GI:47260514
VERSION
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unclassified.
REFERENCE      1 (bases 1 to 13073)
AUTHORS      Belfield,G.P. and Oakley,C.
TITLE      Compositions and methods utilizing the yeast ZEO1 promoter
JOURNAL      Patent: US 6716601-A 20 06-APR-2004;
FEATURES
     source      location/Qualifiers
                1..13073
                /organism="unknown"
                /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 723; DB 6; Length 13073;
Best Local Similarity 100.0%; Pred. No. 6.3e-207;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTCGATTAGCAGCAGACATCATAGACTGGCTGATAAAATATACATACGAAAA 60
Db      16 CTTTCGATTAGCAGCAGACATCATAGACTGGCTGATAAAATATACATACGAAAA 75
Qy      61 ACCATAAAGACAAAGCATCTACTTGTGAAGAAAGAGACGCTTGTGAAGGGGAT 120
Db      76 ACCATAAAGACAAAGCATCTACTTGTGAAGAAAGAGACGCTTGTGAAGGGGAT 135
Qy      121 GGGGGCTAAGAAATCACTTCTTTTCCCTGCGGCTCGGACCGGGGACCCCTCT 180
```

```

Db      136 GGGGGCTAAGAAATCACTTCTTTTCCCTCCGCGTCCGACCCGGAAACCCCTCCT 195
Qy      181 CTCCTCCGACAGATTTCTTCTTTCATATCTTCTTTTATTCCTATCCCGTTGAAGAAC 240
Db      196 CTCCTCCGACAGATTTCTTCTTTCATATCTTCTTTTATTCCTATCCCGTTGAAGAAC 255
Qy      241 GCACTATGACTAATAGTGTCTGGAACATCTCCATAGCTGTGACTTGTGTATCTCAAGT 300
Db      256 GCACTATGACTAATAGTGTCTGGAACATCTCCATAGCTGTGACTTGTGTATCTCAAGT 315
Qy      301 GGTAAACGACCGCTGCTCGAAGACGGTTCCTTGTGACAAATCTAAGAACAGGGGCTACA 360
Db      316 GGTAAACGACCGCTGCTCGAAGACGGTTCCTTGTGACAAATCTAAGAACAGGGGCTACA 375
Qy      361 GTCTCGATATAGAAATATTAAGCGCATTTTGTCTAGCGCGCGCGCGCCGCTTTCCCA 420
Db      376 GTCTCGATATAGAAATATTAAGCGCATTTTGTCTAGCGCGCGCGCGCCGCTTTCCCA 435
Qy      421 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTATTTGGGTAAAGCCCTTTTC 480
Db      436 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTATTTGGGTAAAGCCCTTTTC 495
Qy      481 TGTATTCCGCGCAGTGTGTCTGAGGCTGCGCGCGAGAACATAGTGAATAAGGATGAAC 540
Db      496 TGTATTCCGCGCAGTGTGTGTCTGAGGCTGCGCGCGAGAACATAGTGAATAAGGATGAAC 555
Qy      541 TTTTCATGAGAGATTTAGCAAGCGGAAAAAACTATAGCTAGCTGGAGTTGTTTTTCAA 600
Db      556 TTTTCATGAGAGATTTAGCAAGCGGAAAAAACTATAGCTAGCTGGAGTTGTTTTTCAA 615
Qy      601 TCATATTAAGAGAGAAATTTGTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 660
Db      616 TCATATTAAGAGAGAAATTTGTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 675
Qy      661 TATTGACAGAGACTATCAATCATACAGATATTTGTCAAAAAAAGACTAATATATA 720
Db      676 TATTGACAGAGACTATCAATCATACAGATATTTGTCAAAAAAAGACTAATATATA 735
Qy      721 CAT 723
Db      736 CAT 738

RESULT 4
LOCUS      AR492055      850 bp      DNA      linear      PAT 15-MAY-2004
DEFINITION Sequence 30 from patent US 6716601.
ACCESSION  AR492055
VERSION     AR492055.1  GI:47260524
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 850)
AUTHORS    Belfield,G.P. and Oakley,C.
TITLES     Compositions and methods utilizing the yeast ZEO1 promoter
JOURNAL    Patent: US 6716601-A 30 06-APR-2004;
FEATURES
SOURCE      1..850
             /organism="Unknown"
             /mol_type="genomic DNA"

Query Match      99.6%; Score 720.4; DB 6; Length 850;
Best Local Similarity 99.9%; Pred. No. 2,4e-206;
Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Db      138 ACCATTAAGAGCAAGCAATACCTTGTGAAGAAAGAGACAGCTTTGAAGGGAGT 197
Qy      121 GGGGGCTAAGAAATCACTTCTTTTCCCTGCGGCTCGGACCCGGAAACCCCTCT 180
Db      198 GGGGGCTAAGAAATCACTTCTTTTCCCTGCGGCTCGGACCCGGAAACCCCTCT 257
Qy      181 CTCCTCCGACAGATTTCTTCTTTCATATCTTCTTTTATTCCTATCCCGTTGAAGAAC 240
Db      256 CTCCTCCGACAGATTTCTTCTTTCATATCTTCTTTTATTCCTATCCCGTTGAAGAAC 317
Qy      241 GCACTATGACTAATAGTGTCTGGAACATCTCCATAGCTGTGACTTGTGTATCTCAAGT 300
Db      318 GCACTATGACTAATAGTGTCTGGAACATCTCCATAGCTGTGACTTGTGTATCTCAAGT 377
Qy      301 GGTAAACGACCGCTGCTCGAAGACGGTTCCTTGTGACAAATCTAAGAACAGGGGCTACA 360
Db      378 GGTAAACGACCGCTGCTCGAAGACGGTTCCTTGTGACAAATCTAAGAACAGGGGCTACA 437
Qy      361 GTCTCGATATAGAAATATTAAGCGCATTTTGTCTAGCGCGCGCGCGCCGCTTTCCCA 420
Db      438 GTCTCGATATAGAAATATTAAGCGCATTTTGTCTAGCGCGCGCGCGCCGCTTTCCCA 497
Qy      421 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTTATTTGGGTAAAGCCCTTTTC 480
Db      498 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTTATTTGGGTAAAGCCCTTTTC 557
Qy      481 TGTATTCCGCGCAGTGTGTCTGAGGCTGCGCGCGAGAACATAGTGAATAAGGATGAAC 540
Db      558 TGTATTCCGCGCAGTGTGTGTCTGAGGCTGCGCGCGAGAACATAGTGAATAAGGATGAAC 617
Qy      541 TTTTCATGAGAGATTTAGCAAGCGGAAAAAACTATAGCTAGCTGGAGTTGTTTTTCAA 600
Db      618 TTTTCATGAGAGATTTAGCAAGCGGAAAAAACTATAGCTAGCTGGAGTTGTTTTTCAA 677
Qy      601 TCATATTAAGAGAGAAATTTGTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 660
Db      678 TCATATTAAGAGAGAAATTTGTGCTCACTATGTGACAGTTTCTGGAGCTTTAACTTT 737
Qy      661 TATTGACAGAGACTATCAATCATACAGATATTTGTCAAAAAAAGACTAATATATA 720
Db      738 TATTGACAGAGACTATCAATCATACAGATATTTGTCAAAAAAAGACTAATATATA 797
Qy      721 CA 722
Db      798 AA 799

RESULT 5
LOCUS      SC9920      23498 bp      DNA      linear      PLN 11-AUG-1997
DEFINITION S.cerevisiae chromosome XIII cosmid 9920.
ACCESSION  Z48639.1  GI:732924
VERSION     Z48639.1  GI:732924
KEYWORDS    COX7: cytochrome oxidase; delta element; glutamate decarboxylase;
SOURCE      PETIT; transfer RNA-Ala. (baker's yeast)
ORGANISM    Saccharomyces cerevisiae
REFERENCE   1 (bases 1 to 23498)
AUTHORS    Hunt,S. and Bowman,S.
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 23498)
AUTHORS    Barrell,B., Rajandream,M.A. and Walsh,S.V.
TITLES     Direct Submission
JOURNAL    Submitted (10-MAR-1995) Saccharomyces cerevisiae chromosome XIII
           sequencing project, Sanger Centre, Hinxton Hall, Hinxton, Cambridge
           CB10 1HQ E-mail: barrell@sanger.ac.uk
Notes:
All CDS over 100 codons have been analysed. CDS that are completely
overlapped and those that are overlapped by more than 50%
of their length by a larger CDS have been omitted from this
analysis.

```


Callon, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T.,
 James, K., Eiden, J. A., Rutherford, K., Salzberg, S. L., Craig, A.,
 Kyes, S., Cham, M., Nene, V., Shalimov, S. J., Suh, B., Peterson, J.,
 Angiuoli, S., Perteira, M., Allen, J., Selegut, J., Haft, D.,
 Mather, M. W., Valdeia, A. B., Martin, D. M. A., Faircland, A. H.,
 Frumholz, M. J., Roos, D. S., Ralph, S. A., McPadden, G. I.,
 Cummings, L. M., Subramanian, G. M., Mungall, C., Venter, J. C.,
 Carucci, D. J., Hoffman, S. L., Newbold, C., Davis, R. W., Frazer, C. M. and
 Barrell, B.
 TITLE
 Genome sequence of the human malaria parasite *Plasmodium falciparum*
 PUBMED
 Nature 419 (6906), 498-511 (2002)
 12568864
 2 (bases 1 to 257757)
 Gardner, M. J.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (13-SEP-2002) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 FEATURES
 location/Qualifiers
 1..257757
 /organism="Plasmodium falciparum 3D7"
 /mol_type="genomic DNA"
 /isolate="3D7"
 /db_xref="taxon:36329"
 /chromosome="11"
 repeat region
 257..381

repeat_region	257..381
/rpt_type=tandem	
repeat_region	305..373
/rpt_type=tandem	
repeat_region	763..787
/rpt_type=tandem	
repeat_region	1304..1333
/rpt_type=tandem	
gene	complement(<1434..>.1994)
	/locus_tag="pp11_0070"
mRNA	complement(join(<1434..1828,1910..>.1994))
	/locus_tag="pp11_0070"
CDS	complement(join(1434..1828,1910..1994))
	/locus_tag="pp11_0070"
	/codon_start=1
	/product="hypothetical protein"
	/protein_id="AA035659.1"
	/db_xref="gi:23495997"
	/translation="WCMFVCIIECTBKKKKKKKKLLPYIIHKIDNNKAEONPL
	VVPRIYGHLEVEKRQFPSTLYEVTTILNNDVTKLHAYLVAYERKSLSEFFLGIKK
	ANIVYIKGRKMTKEBNIKSISFSITAVTNSRDEDTISTKVTLFLFLISLTIKV"
repeat_region	1872..1902
/rpt_type=tandem	
repeat_region	1926..1960
/rpt_type=tandem	
repeat_region	2199..2240
/rpt_type=tandem	
repeat_region	2384..2542
/rpt_type=tandem	
repeat_region	2582..2650
/rpt_type=tandem	
gene	complement(<2647..>.4074)
	/locus_tag="pp11_0071"
mRNA	complement(<2647..4074)
	/locus_tag="pp11_0071"
CDS	complement(<2647..4074)
	/locus_tag="pp11_0071"

```

CDS
        /locus_tag="PF11_0071"
        /complement(2647..4074)
        /locus_tag="PF11_0071"
        /codon_start=1
        /product="RuvB DNA helicase, putative"
        /protein_id="AA035660.1"
        /db_xref="GI:2349598"
        /translation="MEINISNITNSNNKKERNVNIHSIKGLGVNTVYILHNEBISLS
        DEKSMFSDNSCGLVGQFAPARESLFLVDLIQKLAGCKLILAGSGSGKSLAIGT
        SRKINRMPVPLSGSEVYSNELKRETVYLEAPRKSIIHKIKEEKLVYEGEVNDWVE
        ENECILYSLNAKOINAIITLKSVKSKTLRLAPRIHQIAEKKIIGVYIYENTNG
        HAVRLRCNDYAEAYDIEPEEYVSLPKGVNKKKEVQOISLHDIILANPVEEDL
        AYSVNLVLPKRIETELKRELVENKVNLESGSAELIPGLYIDEAHMDIIECSEY
        LNAISLAPLAILMATNGICITVGDTGNIPEHGIIPVLDLILIIKPTPYLKEIVQ
        IIALPQTERKINTEPCGVNIAKIVCGSTIPEDLIIPDQV

```



```
repeat_region      QADELFMDAKTSAHRVVDQSNKRFVN"
4159..4258
/rpt_type=tandem
repeat_region      /rpt_type=tandem
4159..4215
/rpt_type=tandem
repeat_region      /rpt_type=tandem
4718..4749
/rpt_type=tandem
repeat_region      /rpt_type=tandem
4922..4978
/rpt_type=tandem
repeat_region      /rpt_type=tandem
5175..5199
/rpt_type=tandem
repeat_region      /rpt_type=tandem
5322..5352
/rpt_type=tandem
repeat_region      /rpt_type=tandem
5350..5397
/rpt_type=tandem
repeat_region      /rpt_type=tandem
5405..5438
/rpt_type=tandem
repeat_region      /rpt_type=tandem
5531..5557
/rpt_type=tandem
gene               /rpt_type=tandem
<5712..>.6914
/locus_tag="PF11_0072"
join(<5712..5802,5896..6191,6311..6454,6636..>.6914)
/locus_tag="PF11_0072"
join(5712..5802,5896..6191,6311..6454,6636..6914)
CDS                /locus_tag="PF11_0072"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAN35661.1"
/db_xref="GI:2349599"
/translation="MKNIKSCVYICLLFPIYIOIMNCLPFGVNIKNKLGVPSPSLIRK
KAVPQPKRQOLIKVNEKGENIFTDGNTNTTKEDTKOTKTNIDIKKIEKEIEI
NLIKSENIILERSFPEKEDITTEILKEITOKYRKIFQKHDKGSSITQII
LFIKIPFLTEHMKNNKDPACLGFLKCVSKRRLLVYIGRDREMEKITVSFNIK
PALPPEYVNDLKYIHFNNTKRFKNAKKKKDKLKKKKVQTDKILFGN"
6250..6288
/rpt_type=tandem
repeat_region      /rpt_type=tandem
6482..6530
/rpt_type=tandem
repeat_region      /rpt_type=tandem
6484..6533
/rpt_type=tandem
repeat_region      /rpt_type=tandem
6913..6938
/rpt_type=tandem
repeat_region      /rpt_type=tandem
7132..7156
/rpt_type=tandem
repeat_region      /rpt_type=tandem
7287..7327
/rpt_type=tandem
repeat_region      /rpt_type=tandem
7691..7726
/rpt_type=tandem
gene               /rpt_type=tandem
<7800..>.10244
/locus_tag="PF11_0073"
join(<7800..9712,9978..10049,10155..10162,10204..>.10244)
/locus_tag="PF11_0073"
join(7800..9712,9978..10049,10155..10162,10204..10244)
CDS                /locus_tag="PF11_0073"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAN35662.1"
/db_xref="GI:2349600"
/translation="MFCSEFLPFFPFVVOONKEIATLGSVHLKKHIIYPRKSHFI
YVVDNNVNLKKNKNEYSKLDKRIKVEYDQKQKIKTELPILPKNNKDYIDH
KKNIDNNNNKKDKKDIINNDKKNKINDTPAQKITNDKKNYNNPEYIENILK
LKNLEVFSSYINKSDVILVPLNYHDI LNKHLEIILPFNNIINSLDYIFLKNNDYK
SININNNNNNNNNNNKYNICPEYNNIYNIIFPIYVNDQDPLNISYTPSPFTKK
DNIIHKEKQLVHIOKPLNDHYNLKDQNEHKYLNINFLTKDNILIDINKIDKDIIE
NNKLYPKKQILNKETNYVNTNTNTNNKQEQDNLKIVHKNDTEFPKKYKDPDLN
LTNENYFIYNPLSSYKLDLKEVIYAKDITINNNAVIGIONINIKKIKHNELLN
NRKCKLANKINSHVAVQELKKQYSYLAHIIISDLLEKDSLLEIKTYLVFVEEN
IQIKTANSIDILFQPKQINHDQLFDYNSKFTLPSNNKYQKQKSYIYYKIKL
LOLQNTANEIINVIYSKGLYKNYSFNEFSFKYTLNVIITTYFNLKNNVNINP
NYLSPNAFGFSSYKDDISLSPKDIMITISEMDQVOKIYIOISLVVPTPKGS"
```

```
repeat_region      /rpt_type=tandem
10060..10092
/rpt_type=tandem
repeat_region      /rpt_type=tandem
10372..10409
/rpt_type=tandem
repeat_region      /rpt_type=tandem
10454..10485
/rpt_type=tandem
gene               /rpt_type=tandem
complement(<10677..>.13503)
/locus_tag="PF11_0074"
complement(join(<10677..10818,10922..10993,11120..11199,
11295..11397,11559..12736,12847..>.13503))
/locus_tag="PF11_0074"
complement(join(10677..10818,10922..10993,11120..11199,
11295..11397,11559..12736,12847..13503))
CDS                /locus_tag="PF11_0074"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAN35663.1"
/db_xref="GI:23496001"
/translation="MIAFLSWVLCTINMAYCISLKKYKFIKSYNLNKKRQONVL
YGIPIRYKWLTSYVPTREBLINNEKQSVDFEYIDMNVIIHCHTANKEKLPYDEH
ELFSNLIQYLKNLFLYIKPEKLIYIGVDGSPKAKNOQKRSFLSIFKINDNTSN
LBNPGCITGTGDFMKINLSKMKPFILKKKVFEDVIFSGSDVAGEEHKILKTYR
ENCKDSDNFKYNNHCICYGLDADLIMSLVTHNNPIFLDKFKISNDLHNNMENIER
TQNTRENDEYNFNPQNDNINYNBGENIQAENKSKTSLVKEBNNTYEENNEIN
NEINNEINNEINNEINNEINNEINNEINDEINKEINNGINNGINNEINNEINNE
INNGINNEINNEINNEINNGNIIIGYKIPFLPVHAYKYEYKYSLNSNDLIDYLR
NIIKTOIATYINKLKKKENIIPVSIIRVVDIYFLSLVGNDFLPHIINIDINEGNE
LINSYIFYIYKSNYITYKDKYHIEBLKILKLSAQEYEPKRGINENISFTDEQ
KRYKYVYIHKPGLDEPKETONIVKRYIEGLFNNLHYHFGCASWYEPYHVALPSD
LISFEKSDPFPEKVKYKKEKLPKGPYSAPTHLSVLPODKNLLDPAYVNIYVEDIVK
SFPENVKIDPDKGKETWYIVHLPINCMINKIITKSKTISKLYLRNKGKQ
IYL"
repeat_region      /rpt_type=tandem
11438..11479
/rpt_type=tandem
repeat_region      /rpt_type=tandem
12185..12400
/rpt_type=tandem
repeat_region      /rpt_type=tandem
12784..12816
/rpt_type=tandem
repeat_region      /rpt_type=tandem
13620..13652
/rpt_type=tandem
repeat_region      /rpt_type=tandem
13757..13788
/rpt_type=tandem
Query Match      6.5%; Score 46.8; DB 3; Length 257757;
Best Local Similarity 53.6%; Pred. No. 0.04;
Matches 120; Conservative 0; Mismatches 102; Indels 2; Gaps 1;
QY      501 TCGAGGCTGCGCGGGAACATPAGATTAAGGATGTAACCTTCGATGAGAGAAATGACA 560
DB      25100 TGAAGATPAGAGAGATTGAACTAATTTAAAGATGACGAGAAAGAAAGAAAGATGA 25159
QY      561 AGCGGAAAAAACTATAGCTAGCTGGAGTTGTTTTCATCATATATAAGGAGAAATT 620
DB      25160 AATGAAAACAATTATGTGATTTAAAGATGTTTCTTATCAACAAATTAACAATT 25219
QY      621 --GTTGCTCACTATGTGACAGTTTCGGGACGCTTAACTTTATTTGACAGAGACATCA 678
DB      25220 TGAATAATCATCTTTTGACGATCGATTGATATATATATATATCTTAAGAAATATCA 25279
QY      679 AATCATACAGATTTTGTCAAAAAAAGAAAGCTATATATACA 722
DB      25280 CGAGAGAACTATATATATGACAGATGAAAGGTAATATATGACA 25323
RESULT 9
BX663508      170295 bp      DNA      linear      VRT 06-JAN-2004
LOCUS      Zebrafish DNA sequence from clone CH211-14486 in linkage group 22,
DEFINITION      complete sequence.
ACCESSION      BX663508
VERSION      BX663508.8      GI:40714027
KEYWORDS      HTG.
```



```

QY          721 CA 722
DB          55928 CA 55927

RESULT 11
AC012052/c 208729 bp DNA linear HTG 04-MAY-2001
LOCUS      Homo sapiens chromosome 2 clone RP11-337F19 map 2, WORKING DRAFT
DEFINITION
AC012052
SEQUENCE   23 unordered pieces.
AC012052
AC012052.3 GI:8576227
HTG; HTGS PHASE1; HTGS _DRAFT.
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 208729)
AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE     Homo sapiens chromosome 2, clone RP11-337F19
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 208729)
AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE     Homo sapiens chromosome 2, clone RP11-337F19
JOURNAL   Unpublished
REFERENCE 1 (bases 1 to 208729)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckert,J.R., Boguslavsky,L., Bouckgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeBrellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lebecky,J., Lien,C., Locke,K., Macdonald,P., Margulis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Tromann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tefaye,S., Tirelli,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (19-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On Jun 21, 2000 this sequence version replaced gi:7321517.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
Center project name: L1385
Center clone name: 337_F_19
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator-amersham; 3% of reads
Chemistry: Dye-terminator Big Dye; 97% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 191830 bases at least Q40
Consensus quality: 199935 bases at least Q30
Consensus quality: 203385 bases at least Q20
Insert size: 210000; agarose-fp
Insert size: 206529; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; agarose-fp
Quality coverage: 4.1 in Q20 bas.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1080: contig of 1080 bp in length

```

```

FEATURES
source
1..208729
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-337F19"
/clone_1b="RPCT-11 Human Male BAC"
1..1080
/note="assembly_fragment"
1181..2268
/note="assembly_fragment"
/note="assembly_fragment"
clone_end:T7
vector_side:right"
2369..3746
/note="assembly_fragment"
3847..6128
/note="assembly_fragment"
6229..7907
/note="assembly_fragment"
8008..10148
/note="assembly_fragment"
10249..12611
/note="assembly_fragment"
12712..14921
/note="assembly_fragment"
15022..18148
/note="assembly_fragment"
1081 1180: gap of 100 bp
1181 2268: contig of 1088 bp in length
2269 2368: gap of 100 bp
2369 3746: contig of 1378 bp in length
3747 3846: gap of 100 bp
3847 6128: contig of 2282 bp in length
6129 6228: gap of 100 bp
6229 7907: contig of 1679 bp in length
7908 8007: gap of 100 bp
8008 10148: contig of 2141 bp in length
10149 10248: gap of 100 bp
10249 12611: contig of 2363 bp in length
12612 12711: gap of 100 bp
12712 14921: contig of 2210 bp in length
14922 15021: gap of 100 bp
15022 18148: contig of 3177 bp in length
18149 18248: gap of 100 bp
18249 22510: contig of 4262 bp in length
22511 22610: gap of 100 bp
22611 27371: contig of 4761 bp in length
27372 27471: gap of 100 bp
27472 32405: contig of 4934 bp in length
32406 32505: gap of 100 bp
32506 38167: contig of 5662 bp in length
38168 38267: gap of 100 bp
38268 46372: contig of 8105 bp in length
46373 46472: gap of 100 bp
46473 5651: contig of 10079 bp in length
5652 5651: gap of 100 bp
5652 64715: contig of 8064 bp in length
64716 64815: gap of 100 bp
64816 75459: contig of 10644 bp in length
75460 75559: gap of 100 bp
75560 90744: contig of 15185 bp in length
90745 90844: gap of 100 bp
90845 107238: contig of 16394 bp in length
107239 107338: gap of 100 bp
107339 128875: contig of 2157 bp in length
128876 128975: gap of 100 bp
128976 152432: contig of 23447 bp in length
152433 152523: gap of 100 bp
152523 178409: contig of 25887 bp in length
178410 208729: gap of 100 bp
208729: contig of 30220 bp in length.
location/Qualifiers
1..208729
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-337F19"
/clone_1b="RPCT-11 Human Male BAC"
1..1080
/note="assembly_fragment"
1181..2268
/note="assembly_fragment"
/note="assembly_fragment"
clone_end:T7
vector_side:right"
2369..3746
/note="assembly_fragment"
3847..6128
/note="assembly_fragment"
6229..7907
/note="assembly_fragment"
8008..10148
/note="assembly_fragment"
10249..12611
/note="assembly_fragment"
12712..14921
/note="assembly_fragment"
15022..18148
/note="assembly_fragment"

```



```

repeat_region complement(5319. .5358)
/rpt_family="AT_rich"
repeat_region complement(5853. .5898)
/rpt_family="AT_rich"
repeat_region complement(5943. .6244)
/rpt_family="MLT1E"
repeat_region complement(6228. .6290)
/rpt_family="MLT1D"
repeat_region complement(6354. .6624)
/rpt_family="L1MA2"
repeat_region complement(6631. .6758)
/rpt_family="MLT1D"
repeat_region complement(7635. .7690)
/rpt_family="(CA)n"
repeat_region 7833. .8023
/rpt_family="MER3"
repeat_region 8188. .8352
/rpt_family="MER5A"
repeat_region 9035. .9057
/rpt_family="AT_rich"
repeat_region 9148. .9227
/rpt_family="MIR"
repeat_region complement(9246. .9302)
/rpt_family="AT_rich"
repeat_region complement(9663. .10244)
/rpt_family="L1MC4"
repeat_region complement(10630. .10701)
/rpt_family="(TA)n"
repeat_region complement(11161. .11220)
/rpt_family="(GA)n"
repeat_region complement(12725. .12846)
/rpt_family="(TA)n"
repeat_region complement(12853. .12932)
/rpt_family="(CAT)n"
repeat_region 13300. .13428
/rpt_family="MIR"
repeat_region complement(13783. .14150)
/rpt_family="L1PA8"
repeat_region 14162. .14188
/rpt_family="AT_rich"
repeat_region complement(14686. .14806)
/rpt_family="FLAM_C"
repeat_region 15736. .15925
/rpt_family="L2"
repeat_region complement(16745. .17110)
/rpt_family="L1M4"
repeat_region 17480. .17941
/rpt_family="MLT1C"
repeat_region 18083. .18134
/rpt_family="MADE1"
repeat_region complement(18177. .18784)
/rpt_family="L1MA9"
repeat_region 19560. .19858
/rpt_family="AluSg"
repeat_region complement(20203. .20229)
/rpt_family="AT_rich"
repeat_region 20350. .20551
/rpt_family="MER4_internal"
repeat_region 20563. .21094
/rpt_family="MER4_internal"
repeat_region 21096. .21246
/rpt_family="MER4A2"
repeat_region 21247. .21510
/rpt_family="MER4A"
repeat_region 21804. .22111
/rpt_family="AluSp"
repeat_region complement(22905. .22958)
/rpt_family="MIR"
repeat_region 23531. .24248
/rpt_family="L1ME3A"
repeat_region complement(24998. .25210)
/rpt_family="L1PA8"
repeat_region 25211. .25359

```

```

repeat_region /rpt_family="MER11A"
25357. .26294
/rpt_family="MER11B"
repeat_region complement(26295. .26395)
/rpt_family="L1PA8"
repeat_region 26608. .26927
/rpt_family="AluSg"
repeat_region 27578. .27878
/rpt_family="AluSx"
repeat_region complement(28196. .28279)
/rpt_family="L12"
repeat_region complement(28709. .28748)
/rpt_family="(TA)n"
repeat_region complement(28749. .29386)
/rpt_family="L1PA4"
STS 29091. .29222
/standard_name="G19948"
/db_xref="dbSTS:32826"
repeat_region 29361. .29554
/rpt_family="L1P"
repeat_region complement(30273. .31278)
/rpt_family="MER11A"
repeat_region complement(32126. .32237)
/rpt_family="(GA)n"
STS 32376. .32676
/standard_name="DXS67"
complement(33176. .33223)
/rpt_family="MIR"
repeat_region 33998. .34051
/rpt_family="MIR"
repeat_region complement(35042. .35071)
/rpt_family="AT_rich"
repeat_region complement(35418. .35439)
/rpt_family="AT_rich"
repeat_region complement(38090. .38123)
/rpt_family="AT_rich"

```

Query Match 5.9%; Score 42.8; DB 9; Length 80659;
Best Local Similarity 51.0%; Pred.No. 0.52;
Matches 101; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```

Qy 517 GAACATAGATGATTAAGGATGATTAACCTTTCATGATGAGATTAAGCGGAAAAAATAT 576
Db 33326 GAACATAGTGTCAAGATTCACAACTCGAGTCAAACTTTCATGATGATTAAGGAAAAATTA 33267
Qy 577 GCGTACGCTGGAGATGTTTTCATCATATTAAGGAGAAATGCTGCTACATATGTA 636
Db 33266 AACATGCTAAGGATTAACCTTAAGTAAATTTTAAGGGAATTAACCTGAGGTTCTGTGC 33207
Qy 637 CAGTTTCTGGACGCTTAACCTTTATATGACAGGACATCAATTCATACAGATATGTC 696
Db 33206 CAGTTTCTCATCAACCAAAAGATATATTAATAAAAAACATTAATCGACATCATTTGTG 33147
Qy 697 AAAAAAAAAAAAAAGACTAA 714
Db 33146 ATGAAAAAAAAAAAAAAA 33129

```

RESULT 13

```

AC021710 195932 bp DNA linear HTG 04-APR-2000
LOCUS Homo sapiens chromosome X clone RP11-12D5 map X, WORKING DRAFT
DEFINITION SEQUENCE, 9 unordered pieces.
AC021710
VERSION AC021710.4 GI:7408016
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 195932)
AUTHORS Birten,B., Linton,L., Nusbaum,C. and Iander,E.
TITLE Homo sapiens chromosome X, clone RP11-12D5

```


together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emr, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/ChrX>

RP6-190D15 is from the library RPCI-6 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pPAC4

IMPORTANT: This sequence is not the entire insert of clone RP6-190D15. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP4-537K23 is at 28759 in this sequence. The true right end of clone RP4-753P9 is at 100 in this sequence.

FEATURES

source

```
1. .28858
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="q25-26.1"
/clone="RP6-190D15"
/clone_1fb="RPCI-6"
39..338
/note="AluY repeat: matches 1..300 of consensus"
347..646
/note="AluSX repeat: matches 1..299 of consensus"
3076..3129
/note="27 copies 2 mer tc 77% conserved"
5914..6224
/note="AluYb8 repeat: matches 1..311 of consensus"
6287..6596
/note="AluSq repeat: matches 1..312 of consensus"
6597..8469
/note="L1MB8 repeat: matches 4267..6133 of consensus"
8470..8757
/note="AluSX repeat: matches 5..302 of consensus"
8758..8796
/note="L1MB8 repeat: matches 6133..6171 of consensus"
8818..9104
/note="AluSg1 repeat: matches 1..289 of consensus"
9345..9465
/note="L2 repeat: matches 2602..2728 of consensus"
10287..10362
/note="38 copies 2 mer aa 69% conserved"
10380..10462
/note="L1MB3 repeat: matches 6035..6127 of consensus"
10464..10518
/note="78K repeat: matches 1..53 of consensus"
10835..11133
/note="AluSX repeat: matches 1..298 of consensus"
11889..12080
/note="AluSg repeat: matches 24..296 of consensus"
12081..12122
/note="AluSg repeat: matches 253..294 of consensus"
12927..12974
/note="24 copies 2 mer gt 72% conserved"
13691..13842
```

```
repeat_region
/note="MIR repeat: matches 65..216 of consensus"
13957..14317
/note="L1M4 repeat: matches 4795..5162 of consensus"
14354..14450
/note="L1M2 repeat: matches 6213..6308 of consensus"
14610..14726
/note="L1M49 repeat: matches 6110..6218 of consensus"
14732..14837
/note="MER8 repeat: matches 49..156 of consensus"
14838..14908
/note="TIGER2 repeat: matches 1..70 of consensus"
14909..15125
/note="MER8 repeat: matches 1..345 of consensus"
15171..15249
/note="TIGER2 repeat: matches 1..73 of consensus"
15250..15315
/note="33 copies 2 mer aa 68% conserved"
15620..15923
/note="AluSq repeat: matches 1..312 of consensus"
15943..15956
/note="L1MB4 repeat: matches 5639..5652 of consensus"
15957..16210
/note="AluSg repeat: matches 38..290 of consensus"
16081..16168
/note="Weakly double-stranded"
16213..16724
/note="L1MB4 repeat: matches 5646..6176 of consensus"
16725..17110
/note="MRT1B repeat: matches 1..388 of consensus"
17111..17277
/note="L1MB4 repeat: matches 5487..5639 of consensus"
17640..17949
/note="AluY repeat: matches 1..309 of consensus"
18445..18745
/note="AluSX repeat: matches 1..302 of consensus"
19047..19348
/note="AluYb repeat: matches 1..291 of consensus"
19523..19744
/note="L1PA12 repeat: matches 5937..6158 of consensus"
19745..20039
/note="L1P4 repeat: matches 5345..5653 of consensus"
20040..20339
/note="AluY repeat: matches 1..301 of consensus"
20391..20689
/note="AluSg repeat: matches 1..299 of consensus"
20699..21369
/note="L1PAB repeat: matches 5489..6163 of consensus"
21430..21459
/note="U6 repeat: matches 1..30 of consensus"
21892..22025
/note="MIR repeat: matches 115..254 of consensus"
23037..23338
/note="AluSX repeat: matches 1..296 of consensus"
23378..23562
/note="L1M4 repeat: matches 5627..5818 of consensus"
23770..23985
/note="AluSX repeat: matches 1..216 of consensus"
23986..24027
/note="Alu repeat: matches 259..298 of consensus"
24570..24787
/note="L1MB repeat: matches 5500..5714 of consensus"
25661..25968
/note="AluSX repeat: matches 1..312 of consensus"
25970..26067
/note="49 copies 2 mer ag 61% conserved"
26552..26865
/note="AluSX repeat: matches 1..312 of consensus"
27019..27089
/note="MERS1B repeat: matches 453..526 of consensus"
27090..27336
/note="AluSC repeat: matches 1..306 of consensus"
27397..27415
/note="MERS1B repeat: matches 525..543 of consensus"
```


repeat_region 27925..28083
/note="12 repeat: matches 2535..2736 of consensus"
repeat_region 28662..28805
/note="MIR repeat: matches 110..256 of consensus"

Query Match 5.8%; Score 42; DB 9; Length 28858;
Best Local Similarity 59.0%; Pred.No. 0.77; Indels 0; Gaps 0;
Matches 72; Conservative 0; Mismatches 50;

QY 590 TTGTTTTCATCATATAAAGGAGAAATTTGCTCATATGACAGTTTCTGGAC 649
DB 13791 TTGTTTCTCATCTGTAAGTGGGAAAGATATCTTTGGCCCAAGTTCCAGGAA 13732
QY 650 GTCTTACTTTATTCGACAGACATCAATCAATATTTGTCACAAAAAAG 709
DB 13731 ATTGTCAGATCAATAGATGTTATGAAAACATTTGAAAAAAGAAAAA 13672
QY 710 AC 711
DB 13671 AC 13670

RESULT 15
AL645468/c 209764 bp DNA linear ROD 26-JUL-2002
LOCUS AL645468
DEFINITION Mouse DNA sequence from clone RP23-246F18 on chromosome 4, complete
ACCESSION AL645468
VERSION AL645468.11 GI:22002669
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 209764)
AUTHORS Blakey, S.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jul 29, 2002 this sequence version replaced gi:21727348.
COMMENT ----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-246F18 is from the RPI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.

FEATURES
source
1..209764
/organism="Mus musculus"

/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-246F18"
/clone_1fb="RPI-23"

Query Match 5.8%; Score 41.8; DB 10; Length 209764;
Best Local Similarity 63.4%; Pred.No. 1.2; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 37;

QY 127 TAAGATGATTCACCTTCTTCCCTTCGCGGTCGACCCGACCCCTCTCC 186
DB 70633 TAAAAATGTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCC 70574
QY 187 GCACGATTTCTTCTTTCATATCTTCTTTATTCCTATCC 227
DB 70573 TTCCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCC 70533

Search completed: February 7, 2005, 22:23:44
Job time : 3606 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2005, 21:09:32 ; Search time 530 Seconds
(without alignments)
8075.421 Million cell updates/sec

Title: US-10-776-213-2

Perfect score: 723

Sequence: 1 cttcgattacgacgcacac.....aaaagactataataacat 723

Scoring table: IDENTITY NUC
Gapop 10-0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1spc04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

. Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	723	100.0	723	4	AAD07475
2	723	100.0	11427	4	AAD07497
3	723	100.0	13073	4	AAD07493
4	720.4	99.6	850	4	AAD07503
5	497.4	68.8	680	6	ABQ76446
6	39	5.4	2000	8	ADA71938
7	38.6	5.3	2311	5	ABSS1325
8	38.4	5.3	544	11	ADR96256
9	37.8	5.2	7736	5	AAS29224
10	37.8	5.2	7736	5	ABA16123
11	37.8	5.2	7736	6	ABG68364
12	37.8	5.2	7736	10	ADC25486
13	37.4	5.2	386	5	ABV04394
14	37	5.1	366	4	AAL35477
15	37	5.1	366	8	ABX58465
16	37	5.1	366	12	ADJ28192
17	37	5.1	390	8	ABX36811
18	37	5.1	400	6	ABO59188
19	36.4	5.0	201	13	ADS41207
20	36.4	5.0	591	13	ACN58226

21	36.4	5.0	110300	13	ADS36499	Adg36499 Human aut
22	36.2	5.0	524	10	ABT22885	ABT22885 Breast ca
23	36	5.0	7624	6	ABL34113	ABL34113 Human imm
24	35.8	5.0	553	13	ACN62578	ACN62578 Cotton de
25	35.8	5.0	815	2	AAZ00806	Aaz00806 Human sec
26	35.8	5.0	815	8	ADA39926	Ada39926 Human sec
27	35.8	5.0	815	8	ACC50510	Acc50510 Human sec
28	35.8	5.0	815	10	ADAS6114	Adas6114 Gene enco
29	35.8	5.0	1143	8	ACA40062	Aca40062 Prokaryot
30	35.6	4.9	10517	13	ADS89438	Ad89438 Oligonuc1
31	35.6	4.9	506	9	ACH17729	Ach17729 Human adu
32	35.6	4.9	5629	6	ABL58962	Ab158962 Human tum
33	35.6	4.9	5988	6	ABK09744	Abk09744 Human ova
34	35.6	4.9	5988	10	ADH29006	Adh29006 Human chr
35	35.6	4.9	5988	12	AD182483	Ad182483 Human mod
36	35.6	4.9	64796	13	ACN37231	Acn37231 Human per
37	35.4	4.9	844	5	ABV18166	Abv18166 Human pro
38	35.4	4.9	33353	4	AAK70003	Aak70003 Human imm
39	35.4	4.9	201239	8	ACA64924	Aca64924 Human PLZ
40	35.2	4.9	281	6	ABV96856	Abv96856 Human pan
41	35.2	4.9	1837	2	AAZ41383	Aaz41383 Human nor
42	35.2	4.9	2000	8	ADA71938	Ada71938 Rice gene
43	35.2	4.9	2938	11	ACN88712	Acn88712 Breast ca
44	35.2	4.9	3197	13	ADS89390	Ad89390 Oligonuc1
45	35.2	4.9	107330	12	ADQ97316	Adq97316 Mouse can

ALIGNMENTS

RESULT 1	
AAD07475	AAD07475 standard; DNA; 723 BP.
XX	XX
AC	AAD07475;
XX	XX
DT	10-AUG-2001 (first entry)
XX	XX
DE	Yeast promoter YMR251WA.
XX	XX
KM	Yeast; promoter; gene expression; fermentable carbon source; glucose;
XX	non-fermentable carbon source; ethanol; yeast cell culture; de.
OS	Saccharomyces cerevisiae.
XX	XX
PN	W0200138549-A1.
XX	XX
PD	31-MAY-2001.
XX	XX
PF	17-NOV-2000; 2000MO-SE002277.
XX	XX
PR	23-NOV-1999; 99SE-00004247.
XX	XX
PA	(ASTR) ASTRAZENECA AB.
XX	XX
PI	Belfield G, Oakley C;
XX	XX
DR	WPI; 2001-367697/38.
XX	XX
PT	New promoter sequences from Saccharomyces cerevisiae useful for
PT	controlling expression of homologous and heterologous nucleic acid
PT	expression in yeast cells.
PS	Claim 1; Page 67-68; 191pp; English.
XX	XX
CC	The invention relates to yeast promoters that are used to control the
CC	expression of homologous and heterologous nucleic acids encoding proteins
CC	and polypeptides in yeast cells. The yeast promoters are induced by a
CC	fermentable carbon source such as glucose or a non-fermentable carbon
CC	source such as ethanol or both. Therefore expression of nucleic acid
CC	molecules encoding a polypeptide under the control of the novel yeast
CC	promoters are regulated by varying the level of a fermentable carbon
CC	source or a non-fermentable carbon source or both. The yeast promoters

CC are useful for, inter alia, the high level production of proteins or
 CC polypeptides in yeast cell culture. The present DNA sequence is
 CC Saccharomyces cerevisiae YMR251AP promoter related to the invention
 XX
 SQ Sequence 723 BP, 199 A; 165 C; 166 G; 193 T; 0 U; 0 Other;

Query Match 100.0%; Score 723; DB 4; Length 723;
 Best Local Similarity 100.0%; Pred. NO. 9.2e-211;

Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CTTTCGATTAGCAGCAGACATCATAGATGCGTCATTAATAATACACTACGAGAAA 60
DB 1 CTTTCGATTAGCAGCAGACATCATAGATGCGTCATTAATAATACACTACGAGAAA 60
QY 61 ACCATAAGAGCAAGGATACCTACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 120
DB 61 ACCATAAGAGCAAGGATACCTACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 120
QY 121 GGGGGCTAAGAGATCATTCATCTTCCCTCGGGGTCGGGAGCCGGGAGCCCTCCT 180
DB 121 GGGGGCTAAGAGATCATTCATCTTCCCTCGGGGTCGGGAGCCGGGAGCCCTCCT 180
QY 181 CTCGCCGACGATTTCTTCTTCATATCTTCTTTATTCCTATCCCGTTGAAGCAACC 240
DB 181 CTCGCCGACGATTTCTTCTTCATATCTTCTTTATTCCTATCCCGTTGAAGCAACC 240
QY 241 GCACATAGACTAAATGGTGTCTGACATCTCATAGGCTGTGACTTGTGTATCTCAGAT 300
DB 241 GCACATAGACTAAATGGTGTCTGACATCTCATAGGCTGTGACTTGTGTATCTCAGAT 300
QY 301 GGTAAACGGACCGGTGCTCGGAAAACGGTCTTCTGTAACAATCTAGAACAGGGGCTACA 360
DB 301 GGTAAACGGACCGGTGCTCGGAAAACGGTCTTCTGTAACAATCTAGAACAGGGGCTACA 360
QY 361 GTCTCGATTAATGAATTAAGAGCGCATTTTGTGAGCGCGCGCGCGCGCTTTCCCA 420
DB 361 GTCTCGATTAATGAATTAAGAGCGCATTTTGTGAGCGCGCGCGCGCGCTTTCCCA 420
QY 421 ATAGAGAGCGCATTTATGAGGAGGCTACTTCTTCTTATTTGGGTAAAGCCCTTTC 480
DB 421 ATAGAGAGCGCATTTATGAGGAGGCTACTTCTTCTTATTTGGGTAAAGCCCTTTC 480
QY 481 TGTTCGCGCAGTGTGCTGACAGGCTGCGCGAGAACATAGTATAGGATGTAAAC 540
DB 481 TGTTCGCGCAGTGTGCTGACAGGCTGCGCGAGAACATAGTATAGGATGTAAAC 540
QY 541 TTTCGATGAGAAATTAAGCAAGCGGAAAAAACTATAGGCTAGCTGGAGGTTGTTTCAA 600
DB 541 TTTCGATGAGAAATTAAGCAAGCGGAAAAAACTATAGGCTAGCTGGAGGTTGTTTCAA 600
QY 601 TCATATTAAGAGGAGAAATTTGCTCACTATGTGACAGTTTCTGGAGGCTCTTAACCTT 660
DB 601 TCATATTAAGAGGAGAAATTTGCTCACTATGTGACAGTTTCTGGAGGCTCTTAACCTT 660
QY 661 TATTCGAGAGGACTATCAATCTATCAAGATTTGTCAAAAAAAGAGCTAATATATA 720
DB 661 TATTCGAGAGGACTATCAATCTATCAAGATTTGTCAAAAAAAGAGCTAATATATA 720
QY 721 CAT 723
DB 721 CAT 723

```

RESULT 2
 AAD07497

ID AAD07497 standard; DNA; 11427 BP.

XX AAD07497;

XX 10-AUG-2001 (first entry)

DE PYMR251AP DNA plasmid.

KW Yeast; promoter; gene expression; fermentable carbon source; glucose;
 KW non-fermentable carbon source; ethanol; yeast cell culture;
 KW PYMR251AP plasmid; ds.

OS Saccharomyces cerevisiae.

OS Unidentified.

OS Chimeric.

PN WC0200138549-A1.

PD 31-MAY-2001.

PF 17-NOV-2000; 2000MO-SE002277.

PR 23-NOV-1999; 99SE-00004247.

PA (ASTR) ASTRAZENECA AB.

PI Belfield G, Oakley C;

DR WPI; 2001-367697/38.

PT New promoter sequences from Saccharomyces cerevisiae useful for
 PT controlling expression of homologous and heterologous nucleic acid
 PT expression in yeast cells.

PS Example 4; Page 133-143; 191BP; English.

CC The invention relates to yeast promoters that are used to control the
 CC expression of homologous and heterologous nucleic acids encoding proteins
 CC and polypeptides in yeast cells. The yeast promoters are induced by a
 CC fermentable carbon source such as glucose or a non-fermentable carbon
 CC source such as ethanol or both. Therefore expression of nucleic acid
 CC molecules encoding a polypeptide under the control of the novel yeast
 CC promoters are regulated by varying the level of a fermentable carbon
 CC source or a non-fermentable carbon source or both. The yeast promoters
 CC are useful for, inter alia, the high level production of proteins or
 CC polypeptides in yeast cell culture. The present sequence is PYMR251AP
 CC plasmid related to the invention. This plasmid contains pBR1 and yeast
 CC YMR251AP promoter

SQ Sequence 11427 BP, 3131 A; 2589 C; 2527 G; 3180 T; 0 U; 0 Other;

Query Match 100.0%; Score 723; DB 4; Length 11427;

Best Local Similarity 100.0%; Pred. NO. 3.3e-210; Indels 0; Gaps 0;

Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CTTTCGATTAGCAGCAGACATCATAGATGCGTCATTAATAATACACTACGAGAAA 60
DB 15 CTTTCGATTAGCAGCAGACATCATAGATGCGTCATTAATAATACACTACGAGAAA 74
QY 61 ACCATAAGAGCAAGGATACCTACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 120
DB 75 ACCATAAGAGCAAGGATACCTACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 134
QY 121 GGGGGCTAAGAGATCATTCATCTTCCCTCGGGTCCGGAGCCGGGAGCCCTCCT 180
DB 135 GGGGGCTAAGAGATCATTCATCTTCCCTCGGGTCCGGAGCCGGGAGCCCTCCT 194
QY 181 CTCGCCGACGATTTCTTCTTCATATCTTCTTTATTCCTATCCCGTTGAAGCAACC 240
DB 195 CTCGCCGACGATTTCTTCTTCATATCTTCTTTATTCCTATCCCGTTGAAGCAACC 254
QY 241 GCACATAGACTAAATGGTGTCTGACATCTCATAGGCTGTGACTTGTGTATCTCAGAT 300
DB 255 GCACATAGACTAAATGGTGTCTGACATCTCATAGGCTGTGACTTGTGTATCTCAGAT 314
QY 301 GGTAAACGGACCGGTGCTCGGAAAACGGTCTTCTGTAACAATCTAGAACAGGGGCTACA 360
DB 315 GGTAAACGGACCGGTGCTCGGAAAACGGTCTTCTGTAACAATCTAGAACAGGGGCTACA 374
QY 361 GTCTCGATTAATGAATTAAGAGCGCATTTTGTGAGCGCGCGCGCGCGCTTTCCCA 420

```

```

Db      375 GTCTCATTAATAGAAATTAAGCCATTTTGTCTAGCGCCGCCGCCGCTTTCCCA 434
Qy      421 ATAGGAGGCGGAGTTTATCGGCGGAGCTCTACTTCTCTATTGGTAAGCCCTTTC 480
Db      435 ATAGGAGGCGGAGTTTATCGGCGGAGCTCTACTTCTCTATTGGTAAGCCCTTTC 494
Qy      481 TGTTCCTGCGGAGTGTCTGCTGAGGCTGCGCGGAGCAATATGTGATAAGGATGAC 540
Db      495 TGTTCCTGCGGAGTGTCTGCTGAGGCTGCGCGGAGCAATATGTGATAAGGATGAC 554
Qy      541 TTTTCATGAGAGATTAAGCAAGCGGAAAAAACTATAGGCTAGCTGGAGTGTTCCTAA 600
Db      555 TTTTCATGAGAGATTAAGCAAGCGGAAAAAACTATAGGCTAGCTGGAGTGTTCCTAA 614
Qy      601 TCATATTAAGGAGAAATTTGTTCTCACTATGTGACAGTTTCTGGAGCTCTTAACCTT 660
Db      615 TCATATTAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCTCTTAACCTT 674
Qy      661 TATTGCAAGAGACTATCAATCTATACATATTTGTCAAAAAAAGACTAATAATA 720
Db      675 TATTGCAAGAGACTATCAATCTATACATATTTGTCAAAAAAAGACTAATAATA 734
Qy      721 CAT 723
Db      735 CAT 737

```

RESULT 3

AAD07493
ID AAD07493 standard; DNA; 13073 BP.

AC AAD07493;

DT 10-AUG-2001 (first entry)

DE pYMR251AP+Luc sequence DNA construct.

XX Yeast; promoter; gene expression; fermentable carbon source; glucose;

KW non-fermentable carbon source; ethanol; yeast cell culture;

KM pYMR251AP plasmid; luciferase gene; de.

XX Saccharomyces cerevisiae.

OS unidentified.

OS Chimeric.

PN WO200138549-A1.

PD 31-MAY-2001.

PF 17-NOV-2000; 2000MO-SE002277.

PR 23-NOV-1999; 99SE-00004247.

XX (ASTR) ASTRAZENECA AB.

PI Belfield G, Oakley C;

DR WPI; 2001-367697/38.

XX New promoter sequences from Saccharomyces cerevisiae useful for

PT controlling expression of homologous and heterologous nucleic acid

XX expression in yeast cells.

XX Example 3; Page 88-99; 191pp; English.

XX The invention relates to yeast promoters that are used to control the

XX expression of homologous and heterologous nucleic acids encoding proteins

XX and polypeptides in yeast cells. The yeast promoters are induced by a

XX fermentable carbon source such as glucose or a non-fermentable carbon

XX source such as ethanol or both. Therefore expression of nucleic acid

XX molecules encoding a polypeptide under the control of the novel yeast

XX promoters are regulated by varying the level of a fermentable carbon

XX source or a non-fermentable carbon source or both. The yeast promoters

CC are useful for, inter alia, the high level production of proteins or

CC polypeptides in yeast cell culture. The present sequence is pYMR251AP+Luc

CC sequence DNA construct related to the invention. The pYMR251AP+Luc

CC sequence contains pPRB1 plasmid, luc backbone (luciferase gene) and yeast

CC YMR251WA promoter

XX

SQ Sequence 13073 BP; 3584 A; 2949 C; 2934 G; 3606 T; 0 U; 0 Other;

Query Match 100.0%; Score 723; DB 4; Length 13073;

Best Local Similarity 100.0%; Pred. No. 3.5e-210;

Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 CTTTCATTAAGCAGCAGACATCTCATAGACTGCTGCTAATAAATACACTACGAAAA 60
Db      16 CTTTCATTAAGCAGCAGACATCTCATAGACTGCTGCTAATAAATACACTACGAAAA 75
Qy      61 ACCATAAAGAGCAAGCAGTACCTACTTGGAAAGAAAGAGAGCAGCTTGTAAAGGGGAT 120
Db      76 ACCATAAAGAGCAAGCAGTACCTACTTGGAAAGAAAGAGAGCAGCTTGTAAAGGGGAT 135
Qy      121 GGGGGCTAAGAAATCAATTCATTTCTTTCCCTTCGCGGTCCGAGCCGGGACCCCTCCT 180
Db      136 GGGGGCTAAGAAATCAATTCATTTCTTTCCCTTCGCGGTCCGAGCCGGGACCCCTCCT 195
Qy      181 CTCCCGGACGATTTCTTCTCTCATATCTTCTTTATTTCCATTCCTGTAAGCAACC 240
Db      196 CTCCCGGACGATTTCTTCTCTCATATCTTCTTTATTTCCATTCCTGTAAGCAACC 255
Qy      241 GCACATAGACTAATGATGCTGGACATCTCAATGCTGATGCTGTGATCTCAAGT 300
Db      256 GCACATAGACTAATGATGCTGGACATCTCAATGCTGATGCTGTGATCTCAAGT 315
Qy      301 GGTAAAGGAGCAGCGTGGCTCGAAAAAGGTTCTTCTGTACAAATTTCTAAGAGGCTACA 360
Db      316 GGTAAAGGAGCAGCGTGGCTCGAAAAAGGTTCTTCTGTACAAATTTCTAAGAGGCTACA 375
Qy      361 GTCTCATTAATAGAAATTAAGAGGCAATTTTGTACGCGCGCGCGCGCGCTTTCCCA 420
Db      376 GTCTCATTAATAGAAATTAAGAGGCAATTTTGTACGCGCGCGCGCGCGCTTTCCCA 435
Qy      421 ATAGGAGGCGGAGTTTATCGGCGGAGCTCTACTTCTTCTATTGGGTAAAGCCCTTTC 480
Db      436 ATAGGAGGCGGAGTTTATCGGCGGAGCTCTACTTCTTCTATTGGGTAAAGCCCTTTC 495
Qy      481 TGTTCCTGCGGAGTGTCTGCTGAGGCTGCGCGGAGCAATATGTATAGGAGTAAAC 540
Db      496 TGTTCCTGCGGAGTGTCTGCTGAGGCTGCGCGGAGCAATATGTATAGGAGTAAAC 555
Qy      541 TTTTCATGAGAAATTAAGCAAGGGAAGAAAAAACTATAGGCTAGCTGGAGTGTTCCTAA 600
Db      556 TTTTCATGAGAAATTAAGCAAGGGAAGAAAAAACTATAGGCTAGCTGGAGTGTTCCTAA 615
Qy      601 TCATATTAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCGTCTTAACCTT 660
Db      616 TCATATTAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCGTCTTAACCTT 675
Qy      661 TATTGCAAGAGACTATCAATCTATACAGATATTTGTCAAAAAAAGACTAATAATA 720
Db      676 TATTGCAAGAGACTATCAATCTATACAGATATTTGTCAAAAAAAGACTAATAATA 725
Qy      721 CAT 723
Db      736 CAT 738

```

RESULT 4

AAD07503
ID AAD07503 standard; DNA; 850 BP.

XX AAD07503;

DT 10-AUG-2001 (first entry)

```

DE  Yeast YMR251WA promoter region.
XX
XX  Yeast; promoter; gene expression; fermentable carbon source; glucose;
KM  non-fermentable carbon source; ethanol; yeast cell culture; ds.
XX
OS  Saccharomyces cerevisiae.
XX
FH  Key
FT  CDS
FT      Location/Qualifiers
FT      1..64
FT      /tag= a
FT      /product= "Yeast YMR251W open reading frame encoding
FT      polypeptide"
FT      complement(78..95)
FT      primer_bind
FT      /tag= C
FT      /bound_molecy= "YMR251WA forward PCR primer"
FT      /partial
FT      786..812
FT      /tag= d
FT      /bound_molecy= "YMR251WA reverse PCR primer"
FT      /partial
FT      CDS
FT      801..850
FT      /tag= b
FT      /product= "Yeast YMR251WA open reading frame encoding
FT      polypeptide"
XX
XX  WO200138549-A1.
XX
XX  31-MAY-2001.
XX
XX  17-NOV-2000; 2000WO-SE002277.
XX
XX  23-NOV-1999; 99SE-00004247.
XX
XX  (ASTR ) ASTRAZENECA AB.
XX
XX  Belfield G, Oakley C;
XX  WPI; 2001-367697/38.
XX
XX  New promoter sequences from Saccharomyces cerevisiae useful for
XX  controlling expression of homologous and heterologous nucleic acid
XX  expression in yeast cells.
XX
XX  Example 3; Fig 14; 191p; English.
XX
XX  The invention relates to yeast promoters that are used to control the
XX  expression of homologous and heterologous nucleic acids encoding proteins
XX  and polypeptides in yeast cells. The yeast promoters are induced by a
XX  fermentable carbon source such as glucose or a non-fermentable carbon
XX  source such as ethanol or both. Therefore expression of nucleic acid
XX  molecules encoding a polypeptide under the control of the novel yeast
XX  promoters are regulated by varying the level of a fermentable carbon
XX  source or a non-fermentable carbon source or both. The yeast promoters
XX  are useful for, inter alia, the high level production of proteins or
XX  polypeptides in yeast cell culture. The present DNA sequence is
XX  Saccharomyces cerevisiae YMR251WA promoter region related to the
XX  invention
XX
XX  Sequence 850 BP; 225 A; 199 C; 193 G; 233 T; 0 U; 0 Other;
SQ
Query Match      99.6%; Score 720.4; DB 4; Length 850;
Best Local Similarity 99.9%; Pred. No. 6.2e-210;
Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY  1 CTTTGATTAGCAGACACATCATGATGCTGCTATAAATATACATACGAAA 60
DB  78 CTTTCATTAGCAGACACATCATGATGCTGCTATAAATATACATACGAAA 137
QY  61 ACCATAAAGAGCAACGATACCTACTTGGAAAGAAAAGAGACGCTTGTAAAGGGGAT 120
DB  138 ACCATAAAGAGCAACGATACCTACTTGGAAAGAAAAGAGACGCTTGTAAAGGGGAT 197
QY  121 GGGGGCTAAGAGTCACTTCACTTTCCCTTCGGCGCTCGGACCCGGACCCCTCT 180

```

```

DB  198 GGGGGCTAAGAGTCACTTCACTTTCCCTTCGGCGCTCGGACCCGGACCCCTCT 257
QY  181 CTTCCCGCAGATTTCTCTCTTTCATATCTCTTTATTTCCATCCGTGAAGCAAC 240
DB  258 CTCCTCCGACGATTTCTCTCTTTCATATCTCTTTATTTCCATCCGTGAAGCAAC 317
QY  241 GCATATGACTAATATGATGCTGACATCTCATAGCTGTGACTTGTGTATCTCAGT 300
DB  318 GCATATGACTAATATGATGCTGACATCTCATAGCTGTGACTTGTGTATCTCAGT 377
QY  301 GGTAAAGGACCGTGGCTCGAAAACGTTCTTCTGTGACAAATTGAAACAGGAGTACA 360
DB  378 GGTAAAGGACCGTGGCTCGAAAACGTTCTTCTGTGACAAATTGAAACAGGAGTACA 437
QY  361 GTCGTATATGAATATATAGGCAATTTTGTGACGCGCGCGCGCGCGCTTCCCA 420
DB  438 GTCTCATATATGAATATATAGGCAATTTTGTGACGCGCGCGCGCGCGCTTCCCA 497
QY  421 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTATTTGGTAAAGCCCTTTC 480
DB  498 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTATTTGGTAAAGCCCTTTC 557
QY  481 TGTTTTCGGCAGTGTGCTGACAGCTCGCGCGGAGAACATATGATTAAGGATGTAAC 540
DB  558 TGTTTTCGGCAGTGTGCTGACAGCTCGCGCGGAGAACATATGATTAAGGATGTAAC 617
QY  541 TTTGATGAGAGATATAGCAACGGAACAACTATAGCTAGCTGGAGTTGTTTTCAA 600
DB  618 TTTGATGAGAGATATAGCAACGGAACAACTATAGCTAGCTGGAGTTGTTTTCAA 677
QY  601 TCATATTAAGGAGAAATTTGCTCACTATGTACAGTTTCTGGACGTTCTTAACTTT 660
DB  678 TCATATTAAGGAGAAATTTGCTCACTATGTACAGTTTCTGGACGTTCTTAACTTT 737
QY  661 TATTGACAGGACTATCAATCATACAGTATATGTCAAAAAAAGACTAATTAATA 720
DB  738 TATTGACAGGACTATCAATCATACAGTATATGTCAAAAAAAGACTAATTAATA 797
QY  721 CA 722
DB  798 AA 799
QY  721 CA 722
DB  798 AA 799
RESULT 5
AB076446
ID  AB076446 standard; cDNA; 680 BP.
XX
XX  AB076446;
XX
XX  21-NOV-2002 (first entry)
XX
XX  S. cerevisiae BAX-associated cDNA fragment SEQ ID 317.
XX
XX  Bax; Bax-resistance; cytosolic; fungicide; immunosuppressive; virucide;
XX  vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
XX  apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
XX  neurodegeneration; cell death; ss.
XX
XX  Saccharomyces cerevisiae.
XX
XX  WO200264766-A2.
XX
XX  22-AUG-2002.
XX
XX  21-DEC-2001; 2001WO-EP015398.
XX
XX  22-DEC-2000; 2000EP-00870318.
XX
XX  04-JAN-2001; 2001EP-00870002.
XX
XX  09-JAN-2001; 2001EP-00870003.
XX
XX  (JANC ) JANSSEN PHARM NV.
XX

```



```

QY 315 GGCTGGAAACGTTCTCTGTCACATTCCTAGACAGGGCTACAGTCTCGATTAATAGA 374
D 424 WKMGCGCMTCRMKSYGMMWRKSKYKMMRSRYRMKRRKCCSTTMGKTRGGMWG 483
QY 375 ATATAAGCCGATTTTCTAGACGCCGCCGCCGCTTCCCATTAAGGAGCCGCA- 433
D 484 TMBGCRFKKSGMKRKRBRMRMRYMRYMSARYMRYCARKKYSASARKRCW 543
QY 434 -GTTTATCGCGAGCTCTACTTTCTTCTATTTGGTAAAGCCCTTCTGTTTGGCCA 492
D 544 YRGGYTAAAMMMKRYRMRYKMMWYKRYSKSYKCMYSYASCMKSARKAGATCK 603
QY 493 GTGTTCTGCAAGCTCGCCGCGAGAACATAGTAAAGGATGTAATCTTCATAGAG 552
D 604 RSKMSAMSKMSRSRKRCKAKSKSSAKRYAMCMGTMSSGMSRMSYTCWRKXGSMKS 663
QY 553 AATTAGCAACGGAAAAAACTATGCTAGCTGGAGTTGTTTTCATCATATAAAGG 612
D 664 TCTMMYMSKYTAKYGSYWRRYRACMWMRYRYRYSYMTYAMWYTSSTMANTGM 723
QY 613 GAGAAATGTTGC 625
D 724 KYSGRYTSMYKY 736

```

RESULT 7

AB551325/c
ID AB551325 standard; cDNA; 2311 BP.

AB551325;

21-OCT-2002 (first entry)

cDNA encoding human secretory protein #23.

Human; secretory polypeptide; SPM; actinic keratosis; arteriosclerosis;
bursitis; cirrhosis; hepatitis; polycythaemia vera; anaemia; psoriasis;
primary thrombocytopenia; cancer; adenocarcinoma; leukaemia; myeloma;
sarcoma; immune system disorder; acquired immunodeficiency syndrome;
AIDS; allergy; asthma; Crohn's disease; diabetes mellitus; gout;
glomerulonephritis; Goodpasture's syndrome; thyroiditis; pancreatitis;
hepatitis; multiple sclerosis; osteoporosis; Reiter's syndrome;
rheumatoid arthritis; neurological disorder; epilepsy; stroke; dementia;
Alzheimer's disease; Pick's disease; Huntington's disease; mood; anxiety;
Parkinson's disease; central nervous system disorder; mental disorder;
schizophrenic disorder; amnesia; Tourette's disorder; transgenic animal;
gene therapy; gene; ss.

Homo sapiens.

WO200257304-A2.

25-JUL-2002.

15-JAN-2002; 2002WO-US001340.

16-JAN-2001; 2001US-0261864P.
16-JAN-2001; 2001US-0261865P.
16-JAN-2001; 2001US-0261979P.
16-JAN-2001; 2001US-0261981P.
17-JAN-2001; 2001US-0262164P.
17-JAN-2001; 2001US-0262208P.
17-JAN-2001; 2001US-0263131P.
19-JAN-2001; 2001US-0262599P.
19-JAN-2001; 2001US-0262760P.
19-JAN-2001; 2001US-0263063P.
19-JAN-2001; 2001US-0263066P.
19-JAN-2001; 2001US-0263070P.
19-JAN-2001; 2001US-0263074P.
19-JAN-2001; 2001US-0263076P.
19-JAN-2001; 2001US-0263077P.
19-JAN-2001; 2001US-0263329P.

(INCY-) INCYTE GENOMICS INC.
Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JT, Jones AL,
Dam TC, Liu TF, Harris B, Flores V, Dafo A, Marwaha R, Chen AJ,
Chang SC, Geretin EH, Petalta CH, David MH, Lewis SA;
WPI; 2002-590716/63.
P-PDB; ABG69833.

New purified secretory polypeptides and polynucleotides. useful in the
diagnosis, study, prevention or treatment of diseases associated with
decreased expression of functional secretory molecules, e.g. AIDS, cancer
or allergies.

Claim 1; Page 260-261; 340p; English.

The invention describes an isolated polynucleotide a naturally occurring
polynucleotide sequence at least 90 % identical to it, a polynucleotide
complementary to it or an RNA equivalent of it. The purified secretory
polypeptides (SPM) and polynucleotides are useful in the diagnosis,
study, prevention or treatment of diseases associated with decreased
expression of functional SPM, e.g. actinic keratosis, arteriosclerosis,
bursitis, cirrhosis, hepatitis, polycythaemia vera, primary
thrombocytopenia, anaemia, psoriasis, cancers including adenocarcinoma,
leukaemia, myeloma or sarcoma, immune system disorder such as acquired
immunodeficiency syndrome (AIDS), allergies, asthma, Crohn's disease,
diabetes mellitus, glomerulonephritis, Goodpasture's syndrome, gout,
Hashimoto's thyroiditis, hepatitis, multiple sclerosis, osteoporosis,
pancreatitis, Reiter's syndrome, autoimmune thyroiditis or rheumatoid
arthritis, neurological disorders such as epilepsy, stroke, Alzheimer's
disease, Pick's disease, Huntington's disease, dementia, Parkinson's
disease, other developmental disorder of the central nervous system,
mental disorder including mood, anxiety or schizophrenic disorder,
amnesia or Tourette's disorder. The polynucleotides may be used in
hybridisation and amplification technologies, e.g. in assessing gene
expression patterns, to develop a transcript image for a particular cell
or tissue, or to create transgenic animals to model human disease. This
sequence encodes a human secretory protein isolated in the invention

Sequence 2311 BP; 824 A; 325 C; 473 G; 689 T; 0 U; 0 Other;

Query Match 5.3%; Score 38.6; DB 6; Length 2311;
Best Local Similarity 53.7%; Pred. No. 0.94;
Matches 80; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 551 AGAATTAGCAACGGAAAAAACTATGCTAGCTGGAGTTGTTTTCATCATATAA 610
D 345 AGTTTAAACAAATTCCTTAATTAATTAATTAATTAATTAATTAATTA 286
QY 611 GGGAGAAATGTTGCTCAGTATGACAGTTCTGGAGCTTTAACTTTATGCAAG 670
D 285 CAGCTTAGTTTCTTAAGTTTAAATTAATTTGGGTGAATTTTAACTTGACAT 226
QY 671 GACTATCAATCATACAGTATTGTCAAA 699
D 225 TAATCATTTTCATGTTTATTTCCAA 197

RESULT 8
ADT96256/c
ID ADT96256 standard; cDNA; 544 BP.

ADT96256;

16-DEC-2004 (first entry)

Colon cancer associated human cDNA sequence #1763.

Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
humoral immune response; cellular immune response; cytoskeletal;
immunostimulant; human; ss.

PR 14-SEP-2000; 2000US-023239P.
 PR 14-SEP-2000; 2000US-023240P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251866P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Baraeh SC, Ruben SM;
 XX
 WPI; 2001-46557/50.
 XX
 PT Nucleic acid molecules encoding human secreted chromosomal binding
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
 XX Alzheimer's and Parkinson's diseases and cancers.
 PS Disclosure; SEQ ID NO 333; 561pp; English.

XX The present invention relates to the isolation of novel DNA-binding
 CC proteins (AAU18154-AAU18281), and cDNA and genomic sequences encoding for
 CC these proteins. DNA-binding, proteins such as histones, chromo (chromatin
 CC organisation modifier) domain proteins, and Y-box binding proteins may
 CC contribute to diseases resulting from aberrant DNA organisation and/or
 CC gene transcription. The sequences of the invention are useful in
 CC screening assays to identify antagonists and/or agonists that may enhance
 CC or block activities mediated by DNA-binding proteins. Blockers of DNA-
 CC binding proteins may be useful in treating disorders such as malignant
 CC diseases (e.g. cancer), autoimmune disorders (e.g. diabetes mellitus),
 CC rheumatic diseases (e.g. rheumatoid arthritis), genetic abnormalities,
 CC (e.g. cystic fibrosis), infectious diseases (e.g. HIV) and neurological
 CC disorders (e.g. Alzheimer's disease). The polynucleotide sequences of the
 CC invention may also be used in gene therapy. AAS29158-AAS29239 represent
 CC genomic sequences encoding for novel DNA-binding proteins. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 7736 BP; 1672 A; 1981 C; 2164 G; 1919 T; 0 U; 0 Other;

Query Match 5.2%; Score 37.8; DB 5; Length 7736;
 Best Local Similarity 54.7%; Pred. No. 2.9;
 Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 110 GTAAGGGGAGTGGGGCTAAGATTCATTCTTTCTTCCCTTCGGGACCG 169
 Db 5184 GTTCAGGGGATGAGTAAAGACACACAGTGTTCCTCCACAGCCGCGAGATGTG 5125
 QY 170 GGACCCCTCTCTCCCGCAGATTTCTCTTCATATCTTCCTTTATTCATCCG 229
 Db 5124 GAAGTACTCTCACTCTCCCGAGCTTCCTTCATATGCTTCGCTGACCTGCTCC 5065
 QY 230 TTGAAGCAACCGCACTA 246
 Db 5064 TGGTAGCAGCTGTACCA 5048

RESULT 10
 ABA16123/C
 ABA16123 standard; DNA; 7736 BP.

XX ABA16123;

XX 23-JAN-2002 (first entry)

DE Human nervous system related polynucleotide SEQ ID NO 8454.

KW Human; nootropic; neuroprotective; cytostatic; dermatological; virocidic;
 KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antischizofrenic; antianaemic; antiallergic; cancer;

KM antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KM antiplatelet; antidiabetic; antifungal; anticonvulsant; antifungal;
KM antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KM neurological disease; infection; nephrotoxic; gene therapy; vaccine; ds.
XX Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001334.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 08-NOV-2000; 2000US-0246615P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249256P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256179P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251859P.
PR 08-DEC-2000; 2000US-0251989P.

PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
PI Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
XX
PS Disclosure; SEQ ID NO 8454; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins
CC (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 7736 BP; 1672 A; 1981 C; 2164 G; 1919 T; 0 U; 0 Other;
XX
Query Match 5.2%; Score 37.6; DB 5; Length 7736;
Best local Similarity 54.7%; Pred. No. 2.9;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
XX
QY 110 GTAAGGGGATGGGGGCTAAGAGTCATTCCTTTCCCTTCGCGGTCCGACCCG 169
Db 5184 GTTCAGGGGATGGATGTAAGACACACAGTGTCCCGCCACGCCCGAGATGTG 5125
QY 170 GGACCCCTCTCTCCCGCAGATTTCTTCTTCAATCTCTTTATTCATCCG 229
Db 5124 GAAGTACTCCATCTCTCCCGAGTCTGCTTCCCTTCATGAGCCTTGACTCC 5065
QY 230 TTGAAGCAGCGCACTA 246
Db 5064 TGGTAGCAGCTTACCA 5048
XX
RESULT 11
ABS68364/c
ID ABS68364 standard; DNA; 7736 BP.
XX
AC ABS68364;
XX
DT 18-NOV-2002 (first entry)
XX
DE Human DNA-binding protein genomic DNA sequence #67.
XX
XX Human; DNA-binding protein; B cell immunodeficiency; autoimmune disorder;
KM severe combined immunodeficiency; rheumatoid arthritis; Crohn's disease;
KM diabetes mellitus; allergy; asthma; inflammatory condition; thrombosis;
KM graft-versus-host disease; blood-related disorder; atherosclerosis;
KM hyperproliferative disorder; cancer; renal disorder; arhythmia;
KM acute glomerulonephritis; cardiovascular disorder; respiratory disorder;
KM Goodpasture's syndrome; neurological disorder; Alzheimer's disease;
KM Parkinson's disease; endocrine disorder; Addison's disease;
KM reproductive system disorder; endometriosis; infectious disease;
KM viral infection; bacterial infection; fungal infection; vaccine;
XX

KW gastrointestinal disorder; multiple sclerosis; gene therapy; ds.
XX Homo sapiens.
XX
XX US2002102638-A1.
XX
PD 01-AUG-2002.
XX
PF 17-JAN-2001; 2001US-00764846.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 20-AUG-2000; 2000US-0226688P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251566P.
PR 08-DEC-2000; 2000US-0251688P.
PR 08-DEC-2000; 2000US-0251689P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX WPI, 2002-690611/74.
XX
XX Novel DNA-binding protein useful for diagnosis, prognosis, prevention and
PT treatment of immune, hyperproliferative, respiratory, cardiovascular,
PT reproductive, endocrine, gastrointestinal and neurological disorders.
XX

PS Claim 1; SEQ ID NO 333; 225bp; English.

XX The present invention relates to a new DNA-binding protein. The invention
CC is useful in treating, preventing, diagnosing and/or prognosing
CC immunodeficiencies (e.g. B cell immunodeficiencies, severe combined
CC immunodeficiencies), autoimmune disorders (rheumatoid arthritis, multiple
CC sclerosis, diabetes mellitus), allergic reactions and conditions (e.g.
CC asthma), inflammatory conditions, graft-versus-host disease, blood-
CC related disorders (thrombosis, atherosclerosis), hyperproliferative
CC disorders (e.g. cancer), renal disorders (e.g. acute glomerulonephritis),
CC cardiovascular disorders (e.g. arrhythmia), respiratory disorders
CC (Goodpasture's syndrome), neurological disorders (e.g. Alzheimer's
CC disease, Parkinson's disease), endocrine disorders (e.g. Addison's
CC disease), reproductive system disorders (e.g. endometriosis) and
CC diseases (e.g. viral, bacterial or fungal infections) and
CC gastrointestinal disorders (e.g. Crohn's disease). The invention is also
CC useful to stimulate neuronal growth and treat, prevent, and/or diagnose
CC neuronal damage which occurs in certain neuronal disorders or neuro-
CC degenerative conditions. The present nucleic acid sequence represents a
CC human DNA-binding protein genomic DNA sequence of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at <http://seqdata.uspto.gov/sequence>

XX

SQ Sequence 7736 BP; 1672 A; 1981 C; 2164 G; 1919 T; 0 U; 0 Other;

Query Match 5.24; Score 37.8; DB 6; Length 7736;
Best Local Similarity 54.74; Pred. No. 2.9;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 110 GTAAGGGGATGGGGGCTAAGACTCATTCATCTTCTTCCCTCCGCTCCGACCCG 169
DB 5184 GTTCAGGGGATGGATGTAAGACACACACAGTGTTCCTCCACACCCCGCAGATGTG 5125
QY 170 GCACCCCTCTCTCCCGACAGATTCTCTTCATATCTCTTATTCATCCG 229
DB 5124 GAAGTACTCCTCTCTCTCCGAGTCTGCTTCTCCCTCATGGCTCTGACTGCTCTCC 5065
QY 230 TTGAAGCAACCGCACTA 246
DB 5064 TGGTAGAGCTGTACCA 5048

RESULT 12
ADC25486/c
ID ADC25486 standard; cDNA; 7736 BP.
XX
AC ADC25486;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human cDNA from extracellular matrix gene 78 #3.
XX
XX Extracellular matrix protein; cytosolic; antibacterial; virulence;
KM neuroprotective; gynaecological; gastrointestinal-Gen; carian;
KM cardiovascular-Gen; nephrotoxic; antiinflammatory; muscular-Gen;
KM respiratory-Gen; immunosuppressive; cerebroprotective; vasodilator;
KM neotrophic; antiallergic; cancer; bacterial infection; viral infection;
KM neural disorder; immune system disorder; blood disorder;
KM muscular disorder; reproductive disorder; gastrointestinal disorder;
KM pulmonary disorder; cardiovascular disorder; renal disorder;
KM inflammatory disorder; proliferative disorder; human; gene therapy; ss;
KM
XX
XX Homo sapiens.
XX
XX US2003049650-A1.
XX
XX 13-MAR-2003.
XX
XX 07-MAR-2002; 2002US-00091483.
XX
XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249256P.
PR 17-NOV-2000; 2000US-0249256P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764846.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX MPI; 2003-605749/57.
DR P-PSDB; ADC25241.
XX
XX
PT New DNA-binding proteins and gene encoding them, useful for diagnosing,
treating and/or preventing e.g. neurological, inflammatory, infectious,

PT Cardiovascular, autoimmune, respiratory, neoplastic or digestive
PT diseases.
XX
XX
XX Disclosure; SEQ ID NO 333; 226pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (cDNA)
CC encoding a human extracellular matrix protein, representing one of 161
CC novel genes. Also included are recombinant vectors, host cells
CC (expressing the protein), the extracellular matrix proteins (including
CC their fragments, epitopes and homologues), an isolated antibody that
CC binds specifically to the protein, diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or absence of a mutation in the nucleic acid and diagnosing a
CC condition based on the presence or absence of the mutation), diagnosing a
CC pathological condition or susceptibility to a pathological condition
CC (comprising determining the presence or amount of expression of the
CC protein in a biological sample and diagnosing a condition based on the
CC presence or amount of expression of the protein), preventing, treating or
CC ameliorating a medical condition by administering the nucleic acid or
CC protein to a mammalian subject, identifying a binding partner to the
CC protein, the gene corresponding to the cDNA sequence, and identifying an
CC activity in a biological assay (comprising expressing the nucleic acid in
CC a cell, isolating the supernatant, detecting an activity in a biological
CC assay and identifying the protein in the supernatant having the
CC activity). The nucleic acids and proteins display the following
CC activities Cytostatic, antibacterial, Virucide, Neuroprotective,

Query Match 5.2%; Score 37.8; DB 10; Length 7736;
Best Local Similarity 54.7%; Pred. No. 2.9;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 110 GTTAAAGGGGATGGGGCTTAAAGTCACTTCTTTCCCTTCCGCGCGGACCG 169
Db 5184 GTTCAGGGGATGGATGTAAAGACACACAGCTGTTCCTCCACAGCCGCGGATG 5125
QY 170 GGACCCCTCTCCCGGACGATTTCTTTCATATCTCTTTATCTTATCCG 229
Db 5124 GAAGTACTCCTCTCTCTCCGAGTCTGCTTCTCCTCCTAGGCTCTGACCTGCTCCC 5065
QY 230 TTGAAGCAACCGCACTA 246
Db 5064 TGTAGCAGCTGTACCA 5048

RESULT 13
ABV04394
ID ABV04394 standard; cDNA; 386 BP.
XX
XX AC ABV04394;
XX
DT 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 4385.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX
XX 23-AUG-2001.
XX
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-02219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 761; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABY00010-ABY62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 386 BP; 108 A; 74 C; 74 G; 128 T; 0 U; 2 Other;

Query Match 5.2%; Score 37.4; DB 5; Length 386;
Best Local Similarity 64.4%; Pred. No. 0.95;
Matches 56; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 628 ACTATGTGACAGTTCTGGACGCTTAATCTTTATTCGAGAGCATCAATCATACA 687
DB 175 ATATATGTGGAACCTCACCTGCTATTTTTCAGGTTCTCTTAAGTAAAGGAAACATCCA 234
OY 688 GATATGTGCAAAAAAAGAGCTAA 714
DB 235 GATACCTTCAAAAAAAGTCTAA 261

RESULT 14
AAL35477/c
ID AAL35477 standard; cDNA; 366 BP.
XX
AC AAL35477;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 819.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiinflammatory; hepatotropic; antidiabetic; antiinflammatory; antitumor;
XX vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein;
XX musculoskeletal system; ss.
XX
XX Homo sapiens.
XX
XX WO200155367-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001338.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231245P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234497P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.

PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218280P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 22-AUG-2000; 2000US-0226668P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI: 2003-128199/12.
DR P-PDB; ABUJ189.
XX
XX
PT Isolated nucleic acid molecules encoding musculoskeletal system
XX associated polypeptides, useful for detecting disorders, e.g. cancer.
XX
XX Claim 1; SEQ ID NO 819; 321pp; English.

CC keratinocyte growth; prevents hair loss, since RGF family members
CC activate hair-forming cells and promotes melanocyte growth; stimulates
CC growth and differentiation of hematopoietic cells and bone marrow cells
CC when used in combination with other cytokines; maintains organs before
CC transplantation or for supporting cell culture of primary tissues;
CC induces tissue of mesodermal origin to differentiate in early embryos;
CC increases or decreases the differentiation or proliferation of embryonic
CC stem cells, besides, haematopoietic lineage; modulates mammalian
CC characteristics, such as, body height, weight, hair colour, eye colour,
CC skin, percentage of adipose tissue, pigmentation, size, and shape (e.g.,
CC cosmetic surgery); modulates mammalian metabolism; changes mammal's metal
CC state or physical state by influencing biorhythms, circadian rhythms,
CC depression, tendency for violence, tolerance for pain, reproductive
CC capabilities, hormonal or endocrine levels, appetite, libido, memory, or
CC stress; increases or decreases storage capabilities, fat content, lipid,
CC protein, carbohydrate, vitamins, minerals, cofactors or other nutritional
CC components. This sequence encodes a novel human musculoskeletal system
CC antigen. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140
XX
SQ Sequence 366 BP; 120 A; 59 C; 53 G; 133 T; 0 U; 1 Other;
Query Match 5.1%; Score 37; DB 8; Length 366;
Best Local Similarity 50.6%; Pred. No. 1.2;
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 549 AGAGATTAGCAAGCGAATAAACTATGCTAGCTGGAGTGTGTTTCAATCATATAA 608
DB 205 AGATTATTTTAAATGAATTAATTAATTAATGCGCTTTAAATGTTGTTAATTTGAAA 146
QY 609 AAGGAGAAATTTGCTCCTATGCTAGTGCAGCTTTCGAGCGCTTAACTTTATTTGCG 668
DB 145 GAAGGACAGTGTCTTCTCTATGCTGCAACCAATCCAAATGTTGGCTTTTAGGTA 86
QY 669 AGACATATCAATCATACGATATTTGCAAAAAAAGCTATATATAA 722
DB 85 AACTCTTCAATTAANACGATTAATACCGTTAAAAAATCCAGATTGATA 32
Search completed: February 7, 2005, 21:23:36
Job time : 534 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2005, 21:13:12 ; Search time 3214 Seconds
(without alignments)
8562.684 Million cell updates/sec

Title: US-10-776-213-2

Perfect score: 723
Sequence: 1 ccttcgatacgcacgcacac.....aaaagacataataacat 723

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g88a1:*
9: gb_g88a2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	249	34.4	434	8	AQ874584 V11B3 mt
C 2	140.8	19.5	927	9	CNS06GVN
C 3	45	6.2	821	9	CNS06GVN
C 4	42.8	5.9	1055	8	CC252581 CH261-136
C 5	42	5.8	1101	9	CNS017KX
C 6	41.6	5.8	1101	9	CNS017KX
C 7	41.4	5.7	436	4	BM274028
C 8	41.4	5.7	1101	9	CNS0100X
C 9	40.8	5.6	791	9	AG532718
C 10	40.4	5.6	698	8	BH572702
C 11	40.4	5.6	791	4	BI255759
C 12	40.2	5.6	1000	9	CNS02UHV
C 13	40	5.5	366	4	BG628729
C 14	40	5.5	518	1	AU077905
C 15	40	5.5	845	8	AZ541311
C 16	40	5.5	871	8	AZ674746
C 17	39.6	5.5	179	2	AM477009
C 18	39.6	5.5	665	6	CA126956
C 19	39.6	5.5	668	9	CE377922
C 20	39.6	5.5	677	9	CL814372
C 21	39.6	5.5	716	9	CL814356
C 22	39.6	5.5	1332	3	CE675673
C 23	39.4	5.4	641	9	CE491438
C 24	39.2	5.4	397	5	BQ399446

C 25	39	5.4	801	7	CK461301
C 26	39	5.4	1233	8	CC192318
C 27	38.8	5.4	932	8	AQ752307
C 28	38.6	5.3	553	7	CN385766
C 29	38.6	5.3	630	8	B19521
C 30	38.6	5.3	677	9	CE498419
C 31	38.6	5.3	1081	9	CL510366
C 32	38.6	5.3	1101	9	CNS00370
C 33	38.4	5.3	243	2	AM653159
C 34	38.4	5.3	594	8	AQ383493
C 35	38.4	5.3	739	9	CE806390
C 36	38.4	5.3	1316	9	AG435289
C 37	38.2	5.3	398	2	AM133335
C 38	38.2	5.3	674	9	AG157962
C 39	38	5.3	600	8	B2312600
C 40	38	5.3	600	8	B2327427
C 41	38	5.3	691	9	CL951462
C 42	38	5.3	802	9	AG468867
C 43	38	5.3	914	8	AZ547087
C 44	38	5.3	931	8	AZ679612
C 45	38	5.3	2639	3	AF289590

ALIGNMENTS

RESULT 1	AQ874584	434 bp	DNA	linear	GSS 08-NOV-1999
LOCUS	V11B3	mtm-3xHA/lacZ	Insertion library, strain Y2278	Saccharomyces	
DEFINITION	cerevisiae genomic 5', genomic survey sequence.				
ACCESSION	AQ874584				
VERSION	AQ874584.1	GI:6286828			
KEYWORDS	GSS.				
SOURCE	Saccharomyces cerevisiae (baker's yeast)				
ORGANISM	Saccharomyces cerevisiae				
REFERENCE	1 (bases 1 to 434)				
AUTHORS	Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., deBerges, S.A., Cheung, K.-H., Sheehan, A., Symoniatie, D., Jansen, R., Umaneky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G.S. and Snyder, M.				
TITLE	Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Kumar A Michael Snyder, Dept. of Mol. Cell. and Dev. Biology Yale University P.O. Box 208103, New Haven, CT 06520-8103, USA Tel: 203 432 9949 Fax: 203 432 6161 Email: anuj.kumare@yale.edu te of mtm-3xHA/lacZ insertion. Seq primer: GGCCTCTTCTTGTGGAAGTAC Class: transposon-tagged Location/Qualifiers 1. 434 /organism="Saccharomyces cerevisiae" /mol_type="genomic DNA" /strain="Y2278 - 5288C background, cir(0) rho(0)" /db_xref="taxon:4932" /lab_host="E. coli" /clone_lib="mtm-3xHA/lacZ insertion library, strain Y2278" /note="Vector: pHS6-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS6-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mtm-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."				
ORIGIN					

Query Match 34.4%; Score 249; DB 8; Length 434;
 Best Local Similarity 100.0%; Pred. No. 6.2e-60;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTGATTAGACGACACACATACATAGACTGCTCATPAAAAATACATAGGAAAA 60
 DB 311 CTTTGATTAGACGACACACATACATAGACTGCTCATPAAAAATACATAGGAAAA 252

QY 61 ACCATTAAGAGAAAGGAGATACCTTCTTGGAAGAAAAGACAGCTTTGTAAGGGGAT 120
 DB 251 ACCATTAAGAGAAAGGAGATACCTTCTTGGAAGAAAAGACAGCTTTGTAAGGGGAT 192

QY 121 GGGGGCTAAGAAGTCACTTCTTCTTCCCTTGCGGGTCCGAGCCCGGGAGCCCTTCT 180
 DB 191 GGGGGCTAAGAAGTCACTTCTTCTTCCCTTGCGGGTCCGAGCCCGGGAGCCCTTCT 132

QY 181 CTTCCCGGACGATTTCTTCTTTCAATCTTCTTTATTCCTATCCGTTGAAGCAACC 240
 DB 131 CTTCCCGGACGATTTCTTCTTTCAATCTTCTTTATTCCTATCCGTTGAAGCAACC 72

QY 241 GCACATAGA 249
 DB 71 GCACATAGA 63

RESULT 2
 CNS06GVN 927 bp DNA linear GSS 30-NOV-2001
 LOCUS T3 end of clone AS0A007G06 of library AS0A from strain CLIB 533
 DEFINITION of Saccharomyces bayanus, genomic survey sequence.
 VERSION AL398217.1 GI:12151528
 KEYWORDS GSS.
 SOURCE Saccharomyces bayanus
 ORGANISM Saccharomyces bayanus
 Eukaryote; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes;
 1 (bases 1 to 927)
 Soulet, J. L., Aigle, M., Artiguenave, F., Blandin, G.,
 de Montigny, J., Dujon, B., Durand, P., Lepoint, A., Llorente, B.,
 Malpertuy, A., Neuvéglise, C., Ozier-Kalogiroupolos, O., Potier, S.,
 Saurin, W., Tekari, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,
 Wincker, P. and Weissenbach, J.
 Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies
 FEBS Lett. 487 (1), 3-12 (2000)
 20584711
 1152876

TITLE 2 (bases 1 to 927)
 JOURNAL Bon, E., Neuvéglise, C., Casaregola, S., Artiguenave, F., Wincker, P.,
 MEDLINE Aigle, M. and Durand, P.
 REFERENCE Genomic exploration of the hemiascomycetous yeasts: 5.
 AUTHORS Saccharomyces bayanus var. uvarum
 JOURNAL FEBS Lett. 487 (1), 37-41 (2000)
 MEDLINE 20584715
 PUBMED 1152880
 3 (bases 1 to 927)
 TITLE Direct Submission
 JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
 MEDLINE 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 REFERENCE segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 JOURNAL This GSS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
 fragilis var. fragilis, Kluyveromyces marxianus var. marxianus, Pichia
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
 5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.

FEATURES
 source location/Qualifiers
 1..927
 /organism="Saccharomyces bayanus"
 /mol_type="genomic DNA"
 /strain="CLIB 533"
 /varietal="uvarum"
 /db_xref="taxon:4931"
 /clone="AS0A007G06"
 /clone_11b="AS0A0A"
 /note="Tend : T3"
 <3..>107
 /note="similar to Saccharomyces cerevisiae ORF YKR251w [strong similarity to YKR076w and YGR154c]"
 /evidence=not_experimental

ORIGIN
 Query Match 19.5%; Score 140.8; DB 9; Length 927;
 Best Local Similarity 69.2%; Pred. No. 6.7e-29;
 Matches 321; Conservative 0; Mismatches 107; Indels 36; Gaps 8;

QY 114 GGGGATAGGGGGCTAAGAATCATTCATTTCTTTCCCTTGGGTCGGAGCCGGGAGC 173
 DB 240 GCGAGAGAGGGGGGTAGGAAGTCACTAC-TCTCTTCCCTTATAGTCCCGAGCGGAAA 298

QY 174 CCTCTCTCCCGGACGATTTCTTCTTCAATCTTCTTTATTCCTATCCGTTGA 233
 DB 299 CCCCCCTTCCCGGACGAGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 345

QY 234 AGCAACCGCATATGCTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 293
 DB 346 GCAACGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 399

QY 294 TCACAGTGTAAAGGACCGGCTCGGAAAGGTTCTTCTGTAACAATTGAAACAG 353
 DB 400 AAACAGCAATAGCGACATCGCGCGCGGAAAC-GTTCTCTAGATGTTGTTAGAACGG 457

QY 354 GGCTACAGTGTGAAT-AGAATAATAAGGCAATTTTGTAGGCGCGCGCGGCGG 406
 DB 458 GGCTATGCGCTGACAAAGTGAATATGAGGCAATTTTGTAGGCGCGCGCGGCGG 517

QY 407 -----GCGCCGCTTCCCAATGAGAGCGGCAATTTTGTAGGCGCGGCGGCGG 461
 DB 518 CCGACCGCGCTTCTTCTTCAATATGAGAGCGGCAATTTTGTAGGCGCGGCGGCGG 576

QY 462 ATTTGGTAAAGCCCTTCTTCTTGTTCGCGCAGTGTCTGAGGCTCGCGGAGAA 521
 DB 577 ATTTGGTAAAGCCCTTCTTCTTGTTCGCGCAGTGTCTGAGGCTCGCGGAGAA 636

QY 522 TAGTATAGGATGTAAAC-TTTCATGAGAGATTTAGCAAGCG 564
 DB 637 GAGCGATTAAGGATGTGACTTTTTCATGAGAGATTTAGCAAGCG 680

RESULT 3
 CNS0090X/C 821 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence TERT3 end of BAC #
 DEFINITION BACR19021 of RPCT-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 VERSION AL052985.1 GI:4934433
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 821)
 TITLE Direct Submission
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)


```

Db      195 AATATTGATTCACCAATTAATTATTCATTAATAATTAATTAATAAACAATGCAAT 136
Qy      713 AATATATACAT 723
Db      135 ATTTATTAATAT 125

RESULT 8
CNS0100X      1101 bp      DNA      linear      GSS 26-JUL-1999
LOCUS      Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION      BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit
ACCESSION      AL098379.1      GI:5609990
VERSION      AL098379
KEYWORDS      GSS.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM      Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
REFERENCE      Direct Submission
AUTHORS      Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
TITLE      BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr
JOURNAL      - Web : www.genoscope.cns.fr
COMMENT      Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billard at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBel0BAC11.

FEATURES
    source
        1..1101
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACN03G04"
            /clone_1ib="DrosBAC"
            /plasmid="pBel0BAC11"
            /note="end : SP6"

ORIGIN
Query Match      5.7%; Score 41.4; DB 9; Length 1101;
Best Local Similarity 20.6%; Pred. No. 2;
Matches 60; Conservative 114; Mismatches 117; Indels 0; Gaps 0;

Qy      433 AGTTATCGCGGAGCTCTACTCTTCCTTATTTGGTAACCCCTTCTGTTTCGCCCA 492
Db      438 WKTDTTGWAGADTDWTDWTTTWTWTGTTTGMKMDTKTKMKWDDDTTDTWTKDW 497
Qy      493 GTGCTGTCGACGCTGCGCGGAGAACATATGATTAAGGATTAACCTTCATGAGAG 552
Db      498 KTDKDDTDKTTDWTANAGMTTDDKTKAKKKRTKTDTKTKTRTGKKKTDGMRKTD 557
Qy      553 AATTACAGCGGAAAAAATGATGCTAGCTGAGGAGTTTTCATCATTAATAAGG 612
Db      558 WGGDRADTDGARDRAANAKODKADGADTDATTDGWTGTTTKADTTDDTDADWDW 617
Qy      613 GAGAAATGTTGCTCACTATGACAGTTTCTGGACGCTTAACTTTATTCAGAGGA 672
Db      618 KDWDMADADTKARKRGKGMGKTKTKTKKKKTDKTKTGTDTKMGKMDTKDKRD 677
Qy      673 CATCAATCATACAGATATTGCAAAAAAAGAGCTAATTAATTAACAT 723
Db      678 GKGGDGRKDKGTGCGDGSRAKGDRAWAMRAKATRAAAAADATAAATAAKAT 728

RESULT 9

```

```

AGS32718
LOCUS      AGS32718      791 bp      DNA      linear      GSS 05-JUN-2004
DEFINITION      Mus musculus molossinus DNA, clone:MSWg01-443E08.TJ, genomic survey
sequence.
ACCESSION      AGS32718
VERSION      AGS32718.1      GI:48293132
KEYWORDS      GSS.
SOURCE      Mus musculus molossinus
ORGANISM      Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE      Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
AUTHORS      BAC end Sequences of Library MSWg01
TITLE      Unpublished
JOURNAL      2 (bases 1 to 791)
REFERENCE      Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
AUTHORS      Direct Submission
TITLE      Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
JOURNAL      and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suohiro-cho,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
[ E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170]
Clones are derived from the mouse BAC library MSWg01. For BAC
library availability, please contact Kunya Abe (abe@rc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9189
e-mail: abe@rc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector : pBAC3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.

FEATURES
    source
        1..791
            /organism="Mus musculus molossinus"
            /mol_type="genomic DNA"
            /sub_species="molossinus"
            /db_xref="taxon:57486"
            /clone="MSWg01-443E08.TJ"
            /sex="male"
            /tissue_type="mixture of kidney and spleen"
            /clone_1ib="MSWg01 Mouse Male BAC library"

ORIGIN
Query Match      5.6%; Score 40.8; DB 9; Length 791;
Best Local Similarity 55.6%; Pred. No. 2.8;
Matches 75; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy      110 GTAAGGGGAGTGGGGGCTAAGAGCATTCACATTCCTTTCCCTGCGGTCGGAACCG 169
Db      290 GAAGCGGAGATGATGGGGGAGACGTTTTTTTCCCNCCCCCCCCCNCCCC 349
Qy      170 GAACCCCTCTCTCCCGGACGATTTCTTCCTTCATATCTTCCTTTATTCCTATCCG 229
Db      350 CCCCCCCTTTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 409
Qy      230 TTGAAGCAACCGAC 244
Db      410 TTTTTCACCTCC 424

RESULT 10
LOCUS      BH572702      698 bp      DNA      linear      GSS 14-DEC-2001
DEFINITION      BGGTD44TF BGGT Brassica oleracea genomic clone BGGTD44, genomic
survey sequence.
ACCESSION      BH572702
VERSION      BH572702.1      GI:17824541
KEYWORDS      GSS.

```

SOURCE	Brassica oleracea
ORGANISM	Brassica oleracea
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE	1 (bases 1 to 698)
AUTHORS	Town,C.D., Van Aken,S., Ultebrack,T., Koo,H. and Fraser,C.M.
TITLE	Whole genome shotgun sequencing of Brassica oleracea
JOURNAL	Unpublished (2001)
COMMENT	Other_GSSs: BCGTD44TR Contact: Chris Town TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TF Class: sheared ends. Location/Qualifiers 1..698 /organism="Brassica oleracea" /mol_type="genomic DNA" /strain="TO100DH3" /db_xref="taxon:3712" /clone="BOGTD44" /clone_1fb="BOGT" /note="Vector: PHOS1; Site 1: BclXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BclXI linkers"
ORIGIN	
Query Match	5.6%; Score 40.4; DB 8; Length 698;
Best Local Similarity	50.0%; Pred. No. 3.5;
Matches	101; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
OY	518 AACATAGGATGATGAGGATGACTTTCATGAGAGAGAAATAGCAACCGCAAAAACATG 577
Db	485 AAAAATTTCTTACTCAATTCGTTTGTGGAGATTTAGTACTGATGATGATGATGATGATG 544
OY	578 GCTAGCTGGAGCTGTTTTCATCATATTAAGGAGAAATGTCGCTACTATGAC 637
Db	545 ACTAAGTTAGATTAATTTTGCACATATATTAGATTTTGGGTTTAATATA 604
OY	638 AGTTCTGGGACGCTTAACTTTTATTCGACAGGACTATCAATATATACAGATATTGCA 697
Db	605 AATTGTTGATGATATATTCGAATTTGATTTAAAAAACAATTTTTTA 664
OY	698 AAAAAAAAAAGACTAATATA 719
Db	665 AAAAAATCACATACTAATATA 686
RESULT 11	
BI255759/c	
LOCUS	BI255759 791 bp mRNA linear EST 17-JUL-2001
DEFINITION	602977180P1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5122410 5',
ACCESSION	BI255759
VERSION	BI255759.1 GI:14809497
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 791)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbe-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Incyte Genomics, Inc.

DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILM at: <http://image.llnl.gov>
 Plate: LLM11298 row: 0 column: 19
 High quality sequence start: 2
 High quality sequence stop: 711.
 Location/Qualifiers
 1..791
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5122410"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

ORIGIN

Query Match	5 6*	Score 40.4;	DB 4;	Length 791;
Beet Local Similarity	58.2*	Pred. No. 3.6;		
Matches	71;	Conservative 0;	Mismatches 51;	Indels 0; Gaps 0;

588 AGTTGTTTTCATCATATAAAGGAGAAATTTGTCCTACTATGTGCAGAGTTTCGGG 647
 505 ACTTCTCCACACATCATAGTCACATGAGAAATTTACAGTGGGAATTCACAAAGTCAGG 446
 648 ACGCTTACTTTTATTTGCAGAGCATATCAATCATACAGATATGTCCAAAAA 707
 445 ATTCTTAATTTTGTGTTGCAGATATCTGAGAGTATAAAGTTTAAAGAAAAA 386
 708 AG 709
 385 TG 384

RESULT 12
 CINS02UHV
 LOCUS
 DEFINITION
 CINS02UHV 1000 bp DNA linear GSS 01-SEP-2000
 167C22 of library G from Tetradon nigroviridis, genomic survey
 sequence.
 AL214492
 AL214492.1 GI:7873311
 GSS: genome survey sequence.
 Tetradon nigroviridis
 Tetradon nigroviridis
 Tetradon nigroviridis
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetradon.

REFERENCE
 1
 Roest Crolius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Pizanes,C., Wincker,P., Broctier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Baseline of human gene number provided by genome-wide analysis
 using Tetradon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)
 20296633
 10835645

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 PUBLISHED
 AUTHORS

2
 Roest Crolius,H., Jalllon,O., Dasilva,C., Ozouf-Costaz,C.,
 Pizanes,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
 Saurin,W., Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetradon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)
 20359837
 10899143
 3 (bases 1 to 1000)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 PUBLISHED
 AUTHORS

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetradon>.
COMMENT Location/Qualifiers
 1..1000
 /organism="Tetradon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone_1ib="G"
 /note="Genoscope sequence ID : COAG167B11LP1-end : T7"
ORIGIN
 Query Match 5.6%; Score 40.2; DB 9; Length 1000;
 Best Local Similarity 45.4%; Pred. No. 4.3;
 Matches 59; Conservative 19; Mismatches 52; Indels 0; Gaps 0;
 QY 91 AAGGAAGAGACAGCTGTAAAGGGGATGGGGCTAGAGATCATCTTCTTTTC 150
 Db 851 AAAAAAAAAAGGGGGGGGGGGGGGGGGGGGAGANNNKKTKKKKBYYYYYY 910
 QY 151 CCTGCGCGTCCGAGACCCGAGCCCTCTCTCCCGAGATTCTCTTCATATCT 210
 Db 911 YYYCCCCCCCCCCCCCCCCCCCCCCCCCCCCCTTTTTTTTTTTTTTTTTT 970
 QY 211 TCCCTTTTATT 220
 Db 971 TTTTTTTTTT 980

RESULT 13
 BG628729/c 366 bp mRNA linear EST 19-APR-2001
LOCUS CC-85f1cLEU23G13d1 Tomato flower library from a mixture of
DEFINITION developmental stages Lycopersicon esculentum cDNA clone
 CC-85f1cLEU23G13d1, mRNA sequence.
ACCESSION BG628729
VERSION BG628729.1 GI:13680202
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 366)
AUTHORS van der Hoeven, R.S. and Tanksey, S.D.
TITLE ESTs from a tomato flower library
JOURNAL Unpublished (2001)
COMMENT Contact: Rutgers S. van der Hoeven
 Cornell University
 252 Emerson Hall, Ithaca, NY 14850, USA
 Tel: 607 255 7886
 Fax: 607 255 6683
 Email: rv19@cornell.edu
 3 prime sequence.
FEATURES Location/Qualifiers
 1..366
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /culivar="E6203"
 /db_xref="taxon:4081"
 /clone="CC-85f1cLEU23G13d1"
 /tissue_type="developing flower buds and open flowers"
 /dev_stage="4-8 week old plants"
 /lab_host="X10LR"
 /clone_1ib="Tomato flower library from a mixture of
 developmental stages"

ORIGIN
 Query Match 5.5%; Score 40; DB 4; Length 366;
 Best Local Similarity 55.9%; Pred. No. 4.1;
 Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
 QY 584 TGGAGTTGTTTTCATCATATAAAGGAGAAATTGTGCTACTATGTGACAGTTTC 643
 Db 136 TGGAGTTGATTTGATTAATTTTATGATTAAGAAGTTTCTTTGGGCTAATAGTTG 77
 QY 644 TGGAGCTCTTAACCTTTATTGACAGAGACTATCAATCATACAGATTTGTAAGAAA 703
 Db 76 CAGGACTTATTTCTTTTGTGCTTATTTGTATCCAGATGAAATTTTCTAAGAAAAA 17
 QY 704 AAAAAAGCTAATATAT 719
 Db 16 AAAAAAAAAAAAAA 1

RESULT 14
 AU077905 518 bp mRNA linear EST 03-APR-2002
LOCUS AU077905 Rice shoot Oryza sativa (japonica cultivar-group) cDNA
DEFINITION clone S5033_82, mRNA sequence.
ACCESSION AU077905
VERSION AU077905.1 GI:5667645
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzoae; Oryza.
REFERENCE 1 (bases 1 to 518)
AUTHORS Yamamoto, K. and Sasaki, T.
TITLE Rice cDNA from etiolated shoot
JOURNAL Unpublished (1997)
COMMENT Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
 305-8602, Japan
 Tel: 81-298-38-7441
 Fax: 81-298-38-7468
 Email: tsasaki@abr.affrc.go.jp, URL: <http://xgp.dna.affrc.go.jp/PROJECT.html>.
FEATURES Location/Qualifiers
 1..518
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /culivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="S5033_82"
 /dev_stage="Etiolated shoot (8 days old)"
 /clone_1ib="Rice shoot"
 /note="Etiolated shoot (8 days old)"
ORIGIN
 Query Match 5.5%; Score 40; DB 1; Length 518;
 Best Local Similarity 57.0%; Pred. No. 4.3;
 Matches 73; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 595 TTTCATCATATATAAAGGAGAAATTGTGCTACTATGTGACAGTTTCTGGACGCTTT 654
 Db 388 TCTGAATGTTAGTGAAGAAAGAGTGAATTCCTGCTGATCAATAAGTGTGT 447
 QY 655 AACTTTATTGACAGAGACTATCAATCATACAGATATTGTCAAAAAAAAAAGACTAA 714
 Db 448 AATTTTGCTCTCAGATGACAAATGAAATCAATCAGAGGTTGATATATAAAAAAAAAA 507
 QY 715 TAATACA 722
 Db 11111

Db 508 AAAAAAA 515

RESULT 15
AZ541311
LOCUS

DEFINITION AZ541311 845 bp DNA linear GSS 14-NOV-2000
ENTDV657R Entamoeba histolytica sheared DNA Entamoeba histolytica
genomic, genomic survey sequence.

ACCESSION AZ541311 GI:11148922
VERSION

KEYWORDS GSS.
Entamoeba histolytica

SOURCE Entamoeba histolytica

ORGANISM Eukaryota; Entamoebidae; Entamoeba.

REFERENCE 1 (bases 1 to 845)

AUTHORS Loftus, B., Van Aken, S. and Fraser, C.
Determination of clone end sequences from Entamoeba histolytica

TITLE HMI:IMSS sheared DNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: b.loftus@tigr.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared

DNA library

Seq primer: M13-Reverse

Class: Shotgun

High quality sequence start: 42

High quality sequence stop: 612.

Location/Qualifiers

1..845

/organism="Entamoeba histolytica"

/mol_type="genomic DNA"

/strain="HMI:IMSS"

/db_xref="taxon:5759"

/clone_id="Entamoeba histolytica sheared DNA"

/note="Vector: pHOSt; Site 1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark,

C.G., and Diamond, E.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.

77:450.). The DNA was mechanically sheared to give a

tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith,

H.O. and Venter, J.C. (Making small insert libraries for

whole genome shotgun sequencing projects. in Genome

Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barell, Oxford University Press, 1999)."

Query Match 5.5%; Score 40; DB 8; Length 845;

Best Local Similarity 55.9%; Pred.No. 4.8; Indels 0; Gaps 0;

Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 588 AGTGTCTTTCATCATATAAGGAGAAATGTTGCTCACTATGTGACAGTTCTGG 647

Db 407 AATTATTTAAATAAAATTCATTTTGAATATTTTATTGTAATGACAGAACTGAA 466

OY 648 ACGTCTTACTTTTATTCAGAGACTATCAATCATACAGATTTGCAAAAAAAA 707

Db 467 AGTTCTTAATTAATAGTAAGAAATATTAATCAATGATGCTTTTCTAACAAAAAT 526

OY 708 AGACTAATTAATACAT 723

Db 527 TTATGAAAAATTAAT 542

Search completed: February 7, 2005, 23:17:23
Job time : 3219 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2005, 21:13:27 ; Search time 176 Seconds

(without alignments)
6721.750 Million cell updates/sec

Title: US-10-776-213-2

Sequence: 1 cttcgattagcagcagcac.....aaaagaccataataacat 723

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 81813859 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	723	100.0	723	US-09-743-194-2	Sequence 2, Appl
2	723	100.0	11427	US-09-743-194-24	Sequence 24, Appl
3	723	100.0	13073	US-09-743-194-20	Sequence 20, Appl
4	720.4	99.6	850	US-09-743-194-30	Sequence 30, Appl
5	47	6.5	7218	US-08-232-463-14	Sequence 14, Appl
6	37.8	5.2	4989	US-09-949-016-17445	Sequence 17445, A
7	37.8	5.2	11440	US-09-949-016-17442	Sequence 17442, A
8	36.6	5.1	16662	US-09-949-016-16850	Sequence 16850, A
9	35.8	5.0	815	US-09-369-247-15	Sequence 15, Appl
10	35.4	4.9	191433	US-09-949-016-16144	Sequence 16144, A
11	35.4	4.9	212449	US-09-949-016-15419	Sequence 15419, A
12	35.2	4.9	72992	US-09-949-016-17592	Sequence 17592, A
13	34.8	4.8	98962	US-09-949-016-14133	Sequence 14133, A
14	34.8	4.8	102884	US-09-949-016-17100	Sequence 17100, A
15	34.4	4.8	601	US-09-949-016-26160	Sequence 26160, A
16	34.4	4.8	601	US-09-949-016-183245	Sequence 183245, A
17	34.4	4.8	511723	US-09-949-016-12152	Sequence 12152, A
18	34.4	4.8	105189	US-09-949-016-16991	Sequence 16991, A
19	34.2	4.7	1876	US-09-469-242-3	Sequence 3, Appl
20	34.2	4.7	14205	US-09-949-016-16196	Sequence 16196, A
21	34.2	4.7	42571	US-09-810-347-3	Sequence 3, Appl
22	34.2	4.7	42571	US-09-949-016-13631	Sequence 13631, A
23	34	4.7	57751	US-09-949-016-13002	Sequence 13002, A
24	34	4.7	232547	US-09-949-016-16603	Sequence 16603, A
25	33.8	4.7	313	US-09-513-999C-15864	Sequence 35864, A
26	33.8	4.7	502	US-09-621-976-18891	Sequence 18891, A
27	33.8	4.7	502	US-09-621-976-18891	Sequence 18891, A

ALIGNMENTS

28	33.8	4.7	4818	3	US-08-817-926-27	Sequence 27, Appl
29	33.6	4.6	77851	4	US-09-949-016-12508	Sequence 12508, A
30	33.6	4.6	77867	4	US-09-949-016-13211	Sequence 13211, A
31	33.6	4.6	77867	4	US-09-949-016-13212	Sequence 13212, A
32	33.6	4.6	77940	4	US-09-949-016-12509	Sequence 12509, A
33	33.4	4.6	2172	4	US-09-976-594-125	Sequence 125, Appl
34	33.4	4.6	312470	4	US-09-949-016-14043	Sequence 14043, A
35	33.4	4.6	336024	4	US-09-949-016-12373	Sequence 12373, A
36	33.2	4.6	828152	4	US-09-949-016-14033	Sequence 14033, A
37	33.2	4.6	828152	4	US-09-949-016-12777	Sequence 12777, A
38	33	4.6	601	4	US-09-949-016-28524	Sequence 28524, A
39	33	4.6	205163	4	US-09-949-016-60957	Sequence 60957, A
40	32.8	4.5	62776	4	US-09-949-016-17576	Sequence 17576, A
41	32.8	4.5	98567	4	US-09-949-016-11750	Sequence 11750, A
42	32.8	4.5	100567	4	US-09-949-016-16934	Sequence 16934, A
43	32.8	4.5	301828	4	US-09-949-016-13969	Sequence 13969, A
44	32.6	4.5	243	4	US-09-248-796A-8463	Sequence 8463, Ap
45	32.6	4.5	243	4	US-09-248-796A-8463	Sequence 8463, Ap

RESULT 1

US-09-743-194-2
Sequence 2, Application US/09743194
Patent No. 671601
GENERAL INFORMATION:
APPLICANT: Belfield, Graham
TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
FILE REFERENCE: 3526.82543
CURRENT FILING DATE: 2001-01-08
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 2
LENGTH: 723
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
US-09-743-194-2

Query Match 100.0%; Score 723; DB 4; Length 723;
Best Local Similarity 100.0%; Pred. No. 1.9e-229;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CTTTGATTAGCAGCAGACATCATAGATGCGTCATATAAATACACTACGGAAG	60
Db	1	CTTTGATTAGCAGCAGACATCATAGATGCGTCATATAAATACACTACGGAAG	60
Qy	61	ACCATTAAGAGCAGCAGATCTCTTCTTGAAGAGAGAGCAGCTTGAAGGGGAT	120
Db	61	ACCATTAAGAGCAGCAGATCTCTTCTTGAAGAGAGAGCAGCTTGAAGGGGAT	120
Qy	121	GGGGGCTTAAGAGTCTTCTTCTTCCCTTCCGCGTCCGAGCCGGAGACCCCTCT	180
Db	121	GGGGGCTTAAGAGTCTTCTTCTTCCCTTCCGCGTCCGAGCCGGAGACCCCTCT	180
Qy	181	CTCCCGCAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	240
Db	181	CTCCCGCAGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	240
Qy	241	GCACTATGATTAATGATGCTGAGCATCTTCATGAGCTGATGATGATCTACAGT	300
Db	241	GCACTATGATTAATGATGCTGAGCATCTTCATGAGCTGATGATGATCTACAGT	300
Qy	301	GCTTAAGGAGCAGCTGCTGGAAGCGTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	360
Db	301	GCTTAAGGAGCAGCTGCTGGAAGCGTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	360
Qy	361	GTTCTGATATGATTAATTAAGGAGCTTTTGTCTAGCGCGCGCGCGCGCTTTCCA	420
Db	361	GTTCTGATATGATTAATTAAGGAGCTTTTGTCTAGCGCGCGCGCGCGCTTTCCA	420

Db	361	GTCTCGATATAATGAATAATTAAGGCAATTTTGTAGGCCGCCGCCGCCCTTTCCCA	420
QY	421	AATAGGAGGCGCAGATTATTCGGCGAGCTCTACTTCTTCCATTTGGGTAAAGCCCTTTC	480
Db	421	AATAGGAGGCGCAGATTATTCGGCGAGCTCTACTTCTTCCATTTGGGTAAAGCCCTTTC	480
QY	481	TGTTTTCCGCCCAAGTGTTGCTGCAGGCTGCGCCGGAGAACATAGTGAATAAGGATGTAAAC	540
Db	481	TGTTTTCCGCCCAAGTGTTGCTGCAGGCTGCGCCGGAGAACATAGTGAATAAGGATGTAAAC	540
QY	541	TTTCGATGAGAAATTAGCAAGCGGAAAAAAACTATGGCTAGCTGGGAGTTGTTTTCAA	600
Db	541	TTTCGATGAGAAATTAGCAAGCGGAAAAAAACTATGGCTAGCTGGGAGTTGTTTTCAA	600
QY	601	TCATATAAAAGGAGAGAAATTGTGCTCATTATGTGACAGTTTCTGGAGCGTCTTAAGCTTT	660
Db	601	TCATATAAAAGGAGAGAAATTGTGCTCATTATGTGACAGTTTCTGGAGCGTCTTAAGCTTT	660
QY	661	TATTCGAGAGAGCTATCAAAATCATACAGATATTTGTCAAAAAAAGAGAGCTATAATATA	720
Db	661	TATTCGAGAGAGCTATCAAAATCATACAGATATTTGTCAAAAAAAGAGAGCTATAATATA	720
QY	721	CAT 723	
Db	721	CAT 723	

```

RESULT 2
US-09-743-194-24
; Sequence 24, Application US/09743194
; Patent No. 6716601
; GENERAL INFORMATION:
; APPLICANT: Oakley, Graham
; APPLICANT: Belfield, Caroline
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
; TITLE OF INVENTION: Controlling Nucleic Acid Expression in Yeast
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/09/743,194
; CURRENT FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 11427
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; US-09-743-194-24

```

Query Match	100.0%	Score 723;	DB 4;	Length 11427;
Best Local Similarity	100.0%	Pred. No. 9,86-229;		
Matches 723; Conservative	0;	Mismatches	0;	Gaps 0;

QY	TTTTGATTTGACAGGCACACATTCACATAGACTCGCTCATTAATAATCACTACGGAAAA	60
Db	CTTTGATTTGACAGGCACACATTCACATAGACTCGCGCATTAATAATCACTACGGAAAA	74
QY	ACCATAAAGACAAAGGATACCTACTTGGAAAGAAAAGAGCA CGCTGTGAAGGGGGAT	120
Db	ACCATAAAGACAAAGGATACCTACTTGGAAAGAAAAGAGCACGCTTGTGAAGGGGGAT	134
QY	GGGGGCTTAAGAAATCATTTCACTTTCTTTTCCCTTCGCGGTCCGGAACCCGGACCCCTCT	180
Db	GGGGGCTTAAGAAAGTCATTTCACTTTCTTTTCCCTTCGCGGTCCGGAACCCGGACCCCTCT	194
QY	CTCCCGCAGATTTCTTCCTTTATATCTTCCTTTATTCCTATCCCGTTGGAAGCAAC	240
Db	CTCCCGCAGATTTCTTCCTTTATATCTTCCTTTATTCCTATCCCGTTGGAAGCAAC	254
QY	GCACTATGACTAAATGGTGTGACATCTCCATGSGCTGTGACTTGTGTATCTCA CAGT	300
Db	GCACTATGACTAAATGGTGTGACATCTCCATGSGCTGTGACTTGTGTATCTCA CAGT	314
QY	GGTAACGGACCCGTTGGCTCGAAACGGTTCTTGTGTACAA TTTGAACAGGGGCTTACA	360

Db	315	GGTAACGGCACCGGTGGCTCGGAAACGGTTCCTTCGTGACAATTTCTGAACACGGGGCTACA	374
Qy	361	GTCTCGATATATAGAAATATATTAAGCGCAATTTTGTCTAGCGCGCGCGCGCCCTTTCCCA	420
Db	375	GTCTCGATATATAGAAATATTAAGCGCATTTTGTCTAGCGCGCGCGGGCGCCGTTTCCCA	434
Qy	421	ATAGGAGAGCGGCAAGTTTATTCGCGCGAGCTTACTTTCTTATTTGGTAAAGCCCTTTC	480
Db	435	ATAGGAGAGCGGCAAGTTTATTCGCGCGAGCTTACTTTCTTATTTGGTAAAGCCCTTTC	494
Qy	481	TGTTTTCCGCGCAGTGGTGTCTCAGAGCTGCGCGCGAGACATAGTAAAGGATGTAC	540
Db	495	TGTTTTCCGCGCAGTGGTGTCTCAGAGCTGCGCGAGAACATAGTATTAAGGATGTAC	554
Qy	541	TTTTCATGAGAGATTAGCAACGGGAAAAAACTATGCTAGCTGGGAGTTGTTTCCA	600
Db	555	TTTTCATGAGAGATTAGCAACGGGAAAAAACTATGCTAGCTGGGAGTTGTTTCCA	614
Qy	601	TCATATTAAGGAGAAATGTGTGCTACTAGTGCACATTTCTGGAGAGTCTTAAGTTT	660
Db	615	TCATATTAAGGAGAAATGTGTGCTACTAGTGCACATTTCTGGAGAGTCTTAAGTTT	674
Qy	661	TATTGACAGAGACTATCAAAATCATACAGATATTTGTCAAAAAAAAAAAAAAGACTAATATAA	720
Db	675	TATTGACAGAGACTATCAAAATCATACAGATATTTGTCAAAAAAAAAAAAAAGACTAATATAA	734
Qy	721	CAT 723	
Db	735	CAT 737	

```

, RESULT 3
, US-09-743-194-20
, Sequence 20, Application US/09743194
, Patent No. 6716601
, GENERAL INFORMATION:
, APPLICANT: Belfield, Graham
, APPLICANT: Oakley, Caroline
, TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
, TITLE OF INVENTION: Controlling Nucleic Acid Expression in Yeast
, FILE REFERENCE: 3526.82343
, CURRENT APPLICATION NUMBER: US/09/743.194
, CURRENT FILING DATE: 2001-01-08
, NUMBER OF SEQ ID NOS: 32
, SOFTWARE: PatentIn Ver. 2.0
, SEQ ID NO:20
, LENGTH: 13073
, TYPE: DNA
, ORGANISM: Saccharomyces cerevisiae
, US-09-743-194-20

```

Query Match	100.0%	Score 723	DB 4	Length 13073
Best Local Similarity	100.0%	Pred. No. 1,1e+28		
Matches 723	0	Mismatches	0	Gaps 0

QY	1	CTTTCAGTTGACAGCACACATCAATGACCTGGCGCTAAAAATACATACGGAGAA	60
Db	16	CTTTCAGTTGACAGCACACATCAATGACCTGGCGCTAAAAATACATACGGAGAA	75
QY	61	ACCTAAAGAGCAAGCGATACCTACTTTGGAGAGAAAGAGACCGCTTGAAGGGGAT	120
Db	76	ACCTAAAGAGCAAGCGATACCTACTTTGGAGAGAAAGAGACCGCTTGAAGGGGAT	135
QY	121	GGGGGCTAAGAGATCATCTTCTTTTCCCTTCGGGCTCGGACCCGGAGCCCTCCT	180
Db	136	GGGGGCTAAGAGATCATCTTCTTTTCCCTTCGGGCTCGGACCCGGAGCCCTCCT	195
QY	181	CTCCCGGACGATTTCTTCTCTTCATATCTTCCCTTTATTCATACCCGTTGAAGCAAC	240
Db	196	CTCCCGGACGATTTCTTCTCTTCATATCTTCTTATTCATACCCGTTGAAGCAAC	255
QY	241	GCACTATGACTAAATGGTGCTGAGACATCTTCATGGCTGTGACTGTGTATCTCAAGT	300

Db	256	GCACATATGACATAAATGGTGGCTGGACATCTCCATGGCTGTGACTTGTGTATCTACAAGT	315
Qy	301	GCTAAACGGCACCCGTGGCTCGGAAAACGGTTCTTTCGTGACAAATTTCTTGAACACAGGGCTTAC	360
Db	316	GGTAACGGCACCGTGGCTCGGAAAACGGTTCTTTCGTGACAAATTTCTTGAACACAGGGCTTAC	375
Qy	361	GTCTCGATTAATATGAATTAATTAAGCCGATTTTTCTAGCGCGCGCGCGGCCCGTTC	420
Db	376	GTCTCGATTAATATGAATTAATTAAGCCGATTTTTCTAGCGCGCGCGCGGCCCGTTC	435
Qy	421	ATAGGAGGCGCAGTTTATTCGGCGGAGCTCTAATTCTTCTTATTTGGTAAAGCCCTTTC	480
Db	436	ATAGGAGGCGCAGTTTATTCGGCGGAGCTCTAATTCTTCTTATTTGGTAAAGCCCTTTC	495
Qy	481	TGTTTTCGGCAGTGTGTGTCTGTCAGAGCTGCGCGCGAGAACATATAGTATTAAGGATGTAC	540
Db	496	TGTTTTCGGCAGTGTGTGTGTCTGTCAGAGCTGCGCGCGAGAACATATAGTATTAAGGATGTAC	555
Qy	541	TTTCGATGAGAAATTTAGCAACGGGAAAAAACTATGCTAGCTGGAGGTTGTTTTCAA	600
Db	556	TTTCGATGAGAAATTTAGCAACGGGAAAAAACTATGCTAGCTGGAGGTTGTTTTCAA	615
Qy	601	TCATATTAATAAGGAGAAATTTGTGTCTCACTATGTGCAACTTTCTGGGAGGTCTTAAC	660
Db	616	TCATATTAATAAGGAGAAATTTGTGTCTCACTATGTGCAACTTTCTGGGAGGTCTTAAC	675
Qy	661	TATTCGAGAGACTATCAAAATCATACAGATATTTGTCAAAAAAAAAAAAAAGACTAATATATA	720
Db	676	TATTCGAGAGACTATCAAAATCATACAGATATTTGTCAAAAAAAAAAAAAAGACTAATATATA	735
Qy	721	CAT 723	
Db	736	CAT 738	

Db	258	CTCCCCGACAGATTTCTTCCTTTCAATCTTCCTTTATTTCTATCCCGGTGAAGACAC	317
Qy	241	GCACATATGACATTAATGCTGCTGGACATCTCCATGGCTGTGACTTGTGTATCTCAACT	300
Db	318	GCACTATATGACATTAATGCTGCTGGACATCTCCATGGCTGTGACTTGTGTATCTCAACT	377
Qy	301	GGTAACGGCACCGGTGGCTCGGAAAAGCTTCCTTCGTGCAAAATTCAGAACAGGGGTACA	360
Db	378	GGTAACGGCACCGGTGGCTCGGAAAAGCTTCCTTCGTGCAAAATTCAGAACAGGGGTACA	437
Qy	361	GTCCTGATTAATAGAATAAATAAGCCGATTTTTGCTATGCGCCGCGCGGCCCGCTTTC	420
Db	438	GTCCTGATTAATAGAATAAATAAGCCGATTTTTGCTATGCGCCGCGCGGCCCGCTTTC	497
Qy	421	ATAGGGAAGCGCAGTTTATTCGGCGGAGCTCACTCTTCCTATTTGGGTAAAGCCCTTTC	480
Db	498	ATAGGGAAGCGCAGTTTATTCGGCGGAGCTCACTCTTCCTATTTGGGTAAAGCCCTTTC	557
Qy	481	TGTTTTGGGCAAGTGCTGCTGACAGGCTGCGCGCGGAGAAACATATGTGATTAAGGATGTAAC	540
Db	558	TGTTTTGGGCAAGTGCTGCTGACAGGCTGCGCGCGGAGAAACATATGTGATTAAGGATGTAAC	617
Qy	541	TTTGGATGAGGAATTTAGACAGCGGAAAAAAACATATGCTATGCTGGGAGTTGTTTTCAA	600
Db	618	TTTGGATGAGGAATTTAGACAGCGGAAAAAAACATATGCTATGCTGGGAGTTGTTTTCAA	677
Qy	601	TCATATTAATAAGGAGAAAATTTGTGCTCATATGTGACAGTTTCTGGAGCGTCTTAAC	660
Db	678	TCATATTAATAAGGAGAAAATTTGTGCTCATATGTGACAGTTTCTGGAGCGTCTTAAC	737
Qy	661	TATTTGCAAGGACATATCAAAATCATACATATATGTCAAAAAAAAGACATTAATAA	720
Db	738	TATTTGCAAGGACATATCAAAATCATACATATATGTCAAAAAAAAGACATTAATAA	797

```

RESULT 4
US-09-743-194-30
; Sequence 30, Application US/09743194
; Patent No. 6716501
; GENERAL INFORMATION:
; APPLICANT: Belfield, Graham
; APPLICANT: Oakley, Caroline
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
; TITLE OF INVENTION: Controlling Nucleic Acid Expression in Yeast
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/09/743,194
; CURRENT FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; US-09-743-194-30

```

Db 798 AA 799

RESULT 5
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHETTLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1600 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA

	Query Match	Similarity	99.6%	Score	720.14	DB	4	Length	850;
	Best Local	Similarity	99.9%	Pred.	1.5e-229				
	Matches	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0
Qy	1	CTTTCGATTACGACCCACACACATACATAGCTGCGTCATTAATAATATACACTACGAAAA							60
Db	78	CTTTCGATTATACACGACACACATACATAGCTGCGTCATTAATAATATACACTACGAAAA							137
Qy	61	ACCATAAAGACCAAGCGATACCTTCTGGAAGAAAAAGAGCACGCTTTGTAAAGGGGAT							120
Db	138	ACCATAAAGACCAAGCGATACCTTCTGGAAGAAAAAGAGCACGCTTTGTAAAGGGGAT							197
Qy	121	GGGGGCTAAGAAGTATTCATCTTTCCTTCCCTCGCGGTCGGGACCGGGAGCCCTCCCT							180
Db	198	GGGGGCTAAGAAGTATTCATCTTTCCTTCCCTCGCGGTCGGGACCGGGAGCCCTCCCT							257
Qy	181	CTCCCGGACGATTTCTTCCTTCAATCTTCTTTTATTCCTATCCGTTGAGAACCC							240

```

1 ZIP: 22313-0299
2
3 COMPUTER READABLE FORM:
4
5 MEDIUM TYPE: Floppy disk
6
7 COMPUTER: IBM PC compatible
8
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10
11 SOFTWARE: PatentIn Release #11.0, Version #1.25
12
13 CURRENT APPLICATION DATA:
14
15 APPLICATION NUMBER: - US/08/232,463
16
17 FILING DATE:
18
19 CLASSIFICATION: 435
20
21 PRIOR APPLICATION DATA:
22
23 APPLICATION NUMBER: US/07/935,313
24
25 FILING DATE:
26
27 APPLICATION NUMBER: EP 91 114 300.6
28
29 FILING DATE: 26-AUG-1991
30
31 ATTORNEY/AGENT INFORMATION:
32
33 NAME: BENT, Stephen A.
34
35 REGISTRATION NUMBER: 29,768
36
37

```

```

? REFERENCE/DOCKET NUMBER: 30472/114 IMM
?
? TELECOMMUNICATION INFORMATION:
?
? TELEPHONE: (703) 836-9300
?
? TELEFAX: (703) 683-4109
?
? TELE: 899149
?
? INFORMATION FOR SEQ ID NO: 14:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 7218 base pairs
?
? TYPE: nucleic acid
?
? STRANDEDNESS: single
?
? TOPOLOGY: linear
?
? IMMEDIATE SOURCE:
?
? CLONE: pTZgpt-Fls
?
? JS-08-232-463-14

```

Query Match 6.5%; Score 47; DB 1; Length 7218;
Best Local Similarity 2.9%; Pred. No. 0.00024;
Matches 11; Conservative 212; Mismatches 152; Indels 0; Gaps 0;

QY	113	AGGGGATGGGGGTAAAGATCATCTTCTTTCCCTCGGGTCGGACCCGGGA	172
Db	1047	AGGTGAGGAGCTTGCAATYYYYYYYYYYYYYYYYYYYYYYYYYYY	1106
QY	173	CCCCCTCTCCC CGCAGATTCTTCTTTCAATCTTCTTTAATCCAGTTG	232
Db	1107	YYY	1166
QY	233	AAGCACCGCACAATGACTAATAATGGTGCAGACATCTCCATGCGCTGACTGTGTGTA	292
Db	1167	YYY	1222
QY	293	CTCA CAGGTGAACGGCACC GNGCTCGGAAA CGSTTCTTGACAA TTTAGA ACAG	352
Db	1227	YYY	1288
QY	353	GAGCTACAGTCTGATAATAGATATAAGCGCA TTTTGCTAGCGCGCCGCCGCC	412
Db	1287	YYY	1344
QY	413	GTTTCCCAATAGGAGGCGCAGTTTATCGGCGAGCTACTTCTTCTAATTTGGGTAA	472
Db	1347	YYY	1400
QY	473	CCCCCTTCTGTTTT C	487
Db	1407	YYYYYYYYYYYYYY 1421	

```

RESULT 6
US-09-949-016-17445
: Sequence 17445, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949, 016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241, 755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237, 768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231, 498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 17445
: LENGTH: 4989
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-17445

```

Query Match	5.2%	Score 37.8;	DB 4;	Length 4989;
Best Local Similarity	54.7%;	Pred. No. 0.22;		
Matches 75; Conservative	0;	Mismatches 62;	Indels 0;	Gaps 0;

<i>OY</i>	170	GACCCCTCCTTCCCAGCATTTCTTCCTTCAATATTCTCTTAATTCATATCCG	229
<i>Db</i>	4524	GAAGTACTTCATCTCTCTCCCGAGTCTGCCTTTCCCTCATGGCCTTGAACTCGGTCCC	4583
<i>OY</i>	230	TTGAAGCAACCGCACTA	246
<i>Db</i>	4584	TGGTAGCAGGTGTACA	4600

```

RESULT 7
US-09-949-016-17442/c
; Sequence 17442, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17442
; LENGTH: 11440
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17442

```

```

Query March 5.2% Score37.8; DB 4; Length 11440;
Best Local Similarity 54.7%; Pred. No. 0.36;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy 110 GTAAGGGGATGGGGGCTAAGAAGTCACTCTTTTCCCTTCGCGATCCGACCG 169
Db 8773 GTTCAGGGGATGGATGTAAAGACACACAGTTGTTCCTCCACAGCCGCCAGATG 8714

Qy 170 GGACCCCTCCCTCTCCCCGACGATTCTTCCCTTGATATCTCTTTATTCCTATCCG 229
Db 8713 GAAGTACTCCACTCTCTCCGAGTGTGCTTCCCTATAGGCGCTGACTGTGCTCC 8654

Qy 230 TTGAAGCAACCGCACTA 246
Db 8653 TGGTAGCAGCTGTACCA 8637

RESULT 8
US-09-949-016-16850/c
; Sequence 16850, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```

```

? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 607/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 16850
? LENGTH: 16662
? TYPE: DNA
? ORGANISM: Human
US-09-949-016-16850

```

Query Match	5.1%;	Score 36.6;	DB 4;	Length 16662;
Best Local Similarity	54.0%;	Pred. No. 1.1;		
Matches	75;	Conservative	0;	Mismatches 64;
			Indels	0;
			Gaps	0;

Oy 89GGAGGAAAAAGAGACCGCTTGTAAAGGGGATGGGGCTTAAGAAGTCATCATTTCTTT 148
Db 4397 GGAAGAAATGAGCAATGCATATAAAGGCTCAGACGCTGCCCTTTTTCTTCCTT 4338

Dy 149 TCCCTTCGCGATCCGGACCCGGAGACCCTCCTTCCCCGACAGTTCTTCTTCATTAT 208
Db 4337 CTTCTTCTTCTTC 4278

Qy	209	CTCCTTTATTCCTATCC	227
Db	4277	CTTCTTTCTTTTCTTTCC	4259

RESULT 9
US-09-369-247-15
Severance 15 Analytication inc/00260247

```

; Patent No. 6569992
;
; GENERAL INFORMATION:
;
; APPLICANT: Rosen et al.
;
; TITLE OF INVENTION: 44 Human Secreted Proteins
;
; PRT SEQUENCE: 270247

```

```

; CURRENT APPLICATION NUMBER: US/09/369,247
; CURRENT FILING DATE: 1999-08-05
; EARLIER APPLICATION NUMBER: 60/074,118
; EARLIER FILING DATE: 1998-02-08

```

EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,137
EARLIER FILING DATE: 1998-02-09

EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,141
EARLIER FILING DATE: 1998-02-09

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 815
;

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (400)

```

```

; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (784)

```

OTHER INFORMATION: n equals a,t,g, or c
US-09-369-247-15

Query Match 5.0%; Score 35.8; DB 4; Length 815;
Best Local Similarity 55.1%; Pred.No. 0.34;
Matches 70; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

554	TAACCTTTAATGCAAGAGCATTAATAAATCCATACAGATATTTCTGCAAAAAAATAAGACTTA	713
594	TTTTCATCATATAAAGGAGAAATTTGTGCTACATATGCAACATTCTGGAGAGCTT	653
665	TATCCTGTTATTTAAATGTGAACATTATTTGACATGAGTGAAGTTATAGTTTAATAGT	724

Db 725 CTGCTCTCTGCGAGCGGTAAATAATATATTTTAAATAAAAA 784
Qy 714 ATTAATTA 720
Db 785 AAAAAA 791

RESULT 10
US-09-949-016-16144
; Sequence 16144, Application US/09949016

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH AND NOT ASSOCIATED WITH

```

```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
;

```

PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; LENGTH: 191433
; TYPE: DNA
; ORGANISM: Human

```

Query Match	4.9%	Score 35.4	DB 4	Length 191433
Best Local Similarity	59.4%	Pred. No. 12		
Matches 60	Conservative 0	Mismatches 41	Indels 0	Gaps 0

QY 104 AGCGTTGTAAGGGGAGTGGGGGCTAGAAGTCATCACTTCTTCCCTTGCGGCTCCG 163

Db 41249 ACGCTGTTAGGGAGAGAGGTAGTAATCAGTGTTCAGATTCCTTTTCTTCAGGTAAAG 41308

Oy 164 GACCCGGGACCCCCTCCTCTCCCCGAGATTTCCTTTC 204
 ||| |||| | ||| | |||
Db 41309 TCCCCCGCACCCCCGGCTGCCCTTGCCCTTCGCAATTTTC 41349

RESULT 11
US-09-949-016-15419/c

Patent No. 681339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016

```

PRIOR APPLICATION NUMBER: 60/241,7555
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,7688
PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ;

```

; SEQ ID NO 15419
; LENGTH: 212449
; TYPE: DNA
; ORGANISM: Homo

```

```
; FEATURE:  
; NAME/KEY: misc_feature  
; LOCATION: (1) .. (212449)  
; OTHER INFORMATION:
```

US-09-949-016-15419

Query Match 4.9%; Score 35.4; DB 4; Length 212449;
Best Local Similarity 57.8%; Pred. No. 13;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 123 GGGCTAAGAGTATCTATCTCTTCTTCCCTTGGCGGTCCGGACCCGGACCCCTCTCT 182
DB 124084 GGTCTCATATTCATTCATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 124025
QY 183 CCCCCGACATTTCT 231
DB 124024 CT 123976

RESULT 12

US-09-949-016-17592/c
; Sequence 17592, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17592
; LENGTH: 72992
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17592

Query Match 4.9%; Score 35.2; DB 4; Length 72992;
Best Local Similarity 60.4%; Pred. No. 7.9;
Matches 58; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 132 AGTATTCACCTTTCTTCCCTTGGCGGTCCGGACCCGGACCCCTCTCTCTCTCTCTCT 191
DB 72622 AGCTCTTCT 72563
QY 192 ATTCT 227
DB 72562 CT 72527

RESULT 13

US-09-949-016-14133/c
; Sequence 14133, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14133

LENGTH: 98962

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(98962)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14133

Query Match 4.8%; Score 34.8; DB 4; Length 98962;
Best Local Similarity 53.7%; Pred. No. 13;
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 84 TACTTGAGAGAAAGAGACAGCTGTAAAGGGGATGGGGGCTAGAGATCTTCACTT 143
DB 78525 TCTGGGAGAGCAGTGGGTGGTGTAGTGGGAGGAGCTTACACCCCAACATTT 78466
QY 144 TCTTTTCCCTTCCGCGGTCCGACCCCGGACCCCTCTCTCCCGCAGATTTCTCTTT 203
DB 78465 TCCCAACCTTGGCCACACACCCCTGGAACCCAGCTCTAGCACTATCTTTT 78406
QY 204 CATATCTTCTTTT 217
DB 78405 TTTTCTTTT 78392

RESULT 14

US-09-949-016-17100/c
; Sequence 17100, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17100
; LENGTH: 102884
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(102884)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17100

Query Match 4.8%; Score 34.8; DB 4; Length 102884;
Best Local Similarity 53.7%; Pred. No. 13;
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 84 TACTTGAGAGAAAGAGACAGCTGTAAAGGGGATGGGGGCTAGAGATCTTCACTT 143
DB 78525 TCTGGGAGAGCAGTGGGTGGTGTAGTGGGAGGAGCTTACACCCCAACATTT 78466
QY 144 TCTTTTCCCTTCCGCGGTCCGACCCCGGACCCCTCTCTCCCGCAGATTTCTCTTT 203
DB 78465 TCCCAACCTTGGCCACACACCCCTGGAACCCAGCTCTAGCACTATCTTTT 78406
QY 204 CATATCTTCTTTT 217
DB 78405 TTTTCTTTT 78392

RESULT 15


```

US-09-949-016-26160/c
: Sequence 26160, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01307
: CURRENT APPLICATION NUMBER: US/09/949, 016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 26160
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-26160

```

Query Match	4.8%	Score 34.4	DB 4	length 601
Best Local Similarity	54.8%	Pred. No. 0.82		
Matches 68; Conservative	0	Mismatches 56	Indels 0	Gaps 0

Qy	25	CAATATGACTCGGTCATAAAAAATACACTACGAAAAACCTAAAGAGCAAGGATCCT	84
Db	409	CAAGAGAGAGAGTCTCAAAAACAAACAAACAAAAAACAAGAGAG	350
Qy	85	ACTTGGAGGAAAAAGAGCAGCGCTTGTAAGGGGATGGGGCTAAGAGTCATTCACCTT	144
Db	349	ACTGGCCAGGAAACAGAGTATGATGAGGCAAAAGAGGTGTGATGAAGAAATCTTATATATGT	290
Qy	145	CTTT	148
Db	289	CATT	286

Search completed: February 7, 2005, 23:20:25
Job time : 180 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 7, 2005, 22:23:48 ; Search time 526 Seconds
(without alignments)
7911.410 Million cell updates/sec

Title: US-10-776-213-2
Perfect score: 1 ctttcgattagcagcagcagc.....aaaagactaataacat 723
Sequence: 723

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4313806 seqs, 2877871033 residues
Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	723	100.0	723	18	US-10-776-213-2
2	723	100.0	11427	18	US-10-776-213-24
3	723	100.0	13073	18	US-10-776-213-20
4	720.4	99.6	850	18	US-10-776-213-30
5	497.4	68.8	580	18	US-10-451-467A-317
6	411	5.7	1024	14	US-10-123-155-198
7	411	5.7	1024	15	US-10-146-731-198
8	411	5.7	1024	15	US-10-140-472-198
9	411	5.7	1024	15	US-10-141-761-198
10	411	5.7	1024	16	US-10-142-885-198
11	411	5.7	1024	16	US-10-158-790-198

C 12	41	5.7	1024	17	US-10-137-871-198	Sequence 198, App
C 13	41	5.7	1024	17	US-10-140-923-198	Sequence 198, App
C 14	41	5.7	1024	17	US-10-141-756-198	Sequence 198, App
C 15	41	5.7	1024	17	US-10-141-759-198	Sequence 198, App
C 16	41	5.7	1024	17	US-10-140-805-198	Sequence 198, App
C 17	41	5.7	1024	17	US-10-140-864-198	Sequence 198, App
C 18	41	5.7	1024	17	US-10-142-426-198	Sequence 198, App
C 19	41	5.7	2561	17	US-10-197-824-20	Sequence 20, App1
C 20	39.4	5.4	409	17	US-10-424-599-16845	Sequence 13, App1
C 21	38.6	5.3	2311	18	US-10-066-531-23	Sequence 23, App1
C 22	38.4	5.3	544	14	US-10-066-543-175	Sequence 175, App
C 23	38.4	5.3	2355	17	US-10-424-599-29921	Sequence 29921, A
C 24	38.2	5.3	1635	17	US-10-424-599-129944	Sequence 129944, A
C 25	37.8	5.2	7736	9	US-09-764-846-333	Sequence 333, App
C 26	37.8	5.2	7736	14	US-10-091-483-333	Sequence 333, App
C 27	37.4	5.2	386	18	US-10-357-930-4385	Sequence 4385, App
C 28	37.2	5.1	436	17	US-10-424-599-96559	Sequence 96559, A
C 29	37.2	5.1	685	13	US-10-027-632-106894	Sequence 106894, A
C 30	37.2	5.1	685	13	US-10-027-632-106894	Sequence 106894, A
C 31	37.2	5.1	685	17	US-10-027-632-106894	Sequence 106894, A
C 32	37.2	5.1	685	17	US-10-027-632-106894	Sequence 106894, A
C 33	37.2	5.1	704	13	US-10-027-632-25182	Sequence 25182, A
C 34	37.2	5.1	704	13	US-10-027-632-25182	Sequence 25182, A
C 35	37.2	5.1	704	17	US-10-027-632-25182	Sequence 25182, A
C 36	37.2	5.1	704	17	US-10-027-632-25182	Sequence 25182, A
C 37	37.2	5.1	1121	17	US-10-424-599-94399	Sequence 94399, A
C 38	37	5.1	366	9	US-09-764-877-819	Sequence 819, App
C 39	37	5.1	366	17	US-10-242-515-819	Sequence 819, App
C 40	37	5.1	390	9	US-09-960-352-1976	Sequence 1976, App
C 41	37	5.1	400	11	US-09-960-034-2883	Sequence 2883, App
C 42	37	5.1	484	18	US-10-425-115-4593	Sequence 4593, App
C 43	36.8	5.1	700	17	US-10-027-632-168936	Sequence 168936, App
C 44	36.8	5.1	700	17	US-10-027-632-168936	Sequence 168936, App
C 45	36.6	5.1	1980090	18	US-10-719-993-6815	Sequence 6815, App

ALIGNMENTS

RESULT 1
US-10-776-213-2
; Sequence 2, Application US/10776213
; Publication NO. US20040142478A1
; GENERAL INFORMATION:
; APPLICANT: AstraZeneca AB
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/10/776, 213
; CURRENT FILING DATE: 2004-02-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; US-10-776-213-2

Query Match 100.0%; Score 723; DB 18; Length 723;
Best Local Similarity 100.0%; Pred. No. 5.5e-217;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CTTTCGATTAGCAGCAGACATCATGACTGGTCAATAAATACTACGGAAGA	60
DB	1	CTTTCGATTAGCAGCAGACATCATGACTGGTCAATAAATACTACGGAAGA	60
QY	61	ACCATTAAGAGCAAGACGATCTTGAAGAAAGAGACAGCCTTGAAGGGGAT	120
DB	61	ACCATTAAGAGCAAGACGATCTTGAAGAAAGAGACAGCCTTGAAGGGGAT	120
QY	121	GGGGGCTAAGAGTCACTTCTTCCCTTCGCGGTCGGAACCCCGGACCCCTCT	180
DB	121	GGGGGCTAAGAGTCACTTCTTCCCTTCGCGGTCGGAACCCCGGACCCCTCT	180

```
QY 181 CTCCCCGACGATTTCTTCTTTGATATCTTCTTTATTCCTAATCCCGTTGAAGCAACC 240
Db 181 CTCCCCGACGATTTCTTCTTTGATATCTTCTTTATTCCTAATCCCGTTGAAGCAACC 240
QY 241 GCACATGACTAATAATGGTGTGACATCTCCATGGCTGTGACTGTGTGTATCTCAAGT 300
Db 241 GCACATGACTAATAATGGTGTGACATCTCCATGGCTGTGACTGTGTGTATCTCAAGT 300
QY 301 GGTAAACGCGACCGTGGCTCGGAAAACGGTTCTTCTGTGACAAATTCAGAACAGGGGCTACA 360
Db 301 GGTAAACGCGACCGTGGCTCGGAAAACGGTTCTTCTGTGACAAATTCAGAACAGGGGCTACA 360
QY 361 GTCGATATATGAAATATATAGCGCATTTTGTGACAGCGCCGCGCGCCGCTTTTCCCA 420
Db 361 GTCGATATATGAAATATATAGCGCATTTTGTGACAGCGCCGCGCGCCGCTTTTCCCA 420
QY 421 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTAATTTGGGTAAAGCCCTTTTC 480
Db 421 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTAATTTGGGTAAAGCCCTTTTC 480
QY 481 TGTTTTGGCGCAGTGTGTCTGACAGGCTCGCGCGGAGAACATAGTAAAGGATGTAAAC 540
Db 481 TGTTTTGGCGCAGTGTGTCTGACAGGCTCGCGCGGAGAACATAGTAAAGGATGTAAAC 540
QY 541 TTTGATGAGAGAAATTAAGCAAGCGGAAAAAACTATGGCTAGCTGGAGTGTTTTCAA 600
Db 541 TTTGATGAGAGAAATTAAGCAAGCGGAAAAAACTATGGCTAGCTGGAGTGTTTTCAA 600
QY 601 TCATATATTAAGGAGAAATTTGTCTCACTATGTGACAGTTTCTGGAGCTCTTAACCTT 660
Db 601 TCATATATTAAGGAGAAATTTGTCTCACTATGTGACAGTTTCTGGAGCTCTTAACCTT 660
QY 661 TATTGCAAGAGACTATCAAAATCATACAGATTTGTCAAAAAAAAGACTAATATTA 720
Db 661 TATTGCAAGAGACTATCAAAATCATACAGATTTGTCAAAAAAAAGACTAATATTA 720
QY 721 CAT 723
Db 721 CAT 723

RESULT 2
US-10-776-213-24
; Sequence 24, Application US/10776213
; Publication No. US20040142478A1
; GENERAL INFORMATION:
; APPLICANT: Astrazeneca AB
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic Acid Expression
; FILE REFERENCE: 3526, 82543
; CURRENT APPLICATION NUMBER: US/10/776, 213
; CURRENT FILING DATE: 2004-02-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 11427
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-776-213-24

Query Match 100.0%; Score 723; DB 18; Length 11427;
Best Local Similarity 100.0%; Pred. No. 2, 5e-216;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CTTTCGATTAAGCAAGCAATCACTTGGAGAGAAAAGAGACGCTTGTAAAGGGGAT 120
Db 15 CTTTCGATTAAGCAAGCAATCACTTGGAGAGAAAAGAGACGCTTGTAAAGGGGAT 120
```

```
QY 61 ACCATAAAGAGCAAGCAATCACTTGGAGAGAAAAGAGACGCTTGTAAAGGGGAT 120
Db 75 ACCATAAAGAGCAAGCAATCACTTGGAGAGAAAAGAGACGCTTGTAAAGGGGAT 134
```

```
QY 121 GGGGGCTAAGAGATCACTTCTTTTCCCTTCGCGGTCCGGACCCGGACCCCTCCT 180
Db 135 GGGGGCTAAGAGATCACTTCTTTTCCCTTCGCGGTCCGGACCCGGACCCCTCCT 194
QY 181 CTCCCCGACGATTTCTTCTTTGATATCTTCTTTATTCCTAATCCCGTTGAAGCAACC 240
Db 195 CTCCCCGACGATTTCTTCTTTGATATCTTCTTTATTCCTAATCCCGTTGAAGCAACC 254
QY 241 GCACATGACTAATAATGGTGTGACATCTCCATGGCTGTGACTGTGTGTATCTCAAGT 300
Db 255 GCACATGACTAATAATGGTGTGACATCTCCATGGCTGTGACTGTGTGTATCTCAAGT 314
QY 301 GGTAAACGCGACCGTGGCTCGGAAAACGGTTCTTCTGTGACAAATTCAGAACAGGGGCTACA 360
Db 315 GGTAAACGCGACCGTGGCTCGGAAAACGGTTCTTCTGTGACAAATTCAGAACAGGGGCTACA 374
QY 361 GTCGATATATGAAATATATAGCGCATTTTGTGACAGCGCCGCGCGCCGCTTTTCCCA 420
Db 375 GTCGATATATGAAATATATAGCGCATTTTGTGACAGCGCCGCGCGCCGCTTTTCCCA 434
QY 421 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTAATTTGGGTAAAGCCCTTTTC 480
Db 435 ATAGGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTAATTTGGGTAAAGCCCTTTTC 494
QY 481 TGTTTTGGCGCAGTGTGTCTGACAGGCTCGCGCGGAGAACATAGTAAAGGATGTAAAC 540
Db 495 TGTTTTGGCGCAGTGTGTCTGACAGGCTCGCGCGGAGAACATAGTAAAGGATGTAAAC 554
QY 541 TTTGATGAGAGAAATTAAGCAAGCGGAAAAAACTATGGCTAGCTGGAGTGTTTTCAA 600
Db 555 TTTGATGAGAGAAATTAAGCAAGCGGAAAAAACTATGGCTAGCTGGAGTGTTTTCAA 614
QY 601 TCATATATTAAGGAGAAATTTGTCTCACTATGTGACAGTTTCTGGAGCTCTTAACCTT 660
Db 615 TCATATATTAAGGAGAAATTTGTCTCACTATGTGACAGTTTCTGGAGCTCTTAACCTT 674
QY 661 TATTGCAAGAGACTATCAAAATCATACAGATTTGTCAAAAAAAAGACTAATATTA 720
Db 675 TATTGCAAGAGACTATCAAAATCATACAGATTTGTCAAAAAAAAGACTAATATTA 734
QY 721 CAT 723
Db 735 CAT 737
```

```
RESULT 3
US-10-776-213-20
; Sequence 20, Application US/10776213
; Publication No. US20040142478A1
; GENERAL INFORMATION:
; APPLICANT: Astrazeneca AB
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic Acid Expression
; FILE REFERENCE: 3526, 82543
; CURRENT APPLICATION NUMBER: US/10/776, 213
; CURRENT FILING DATE: 2004-02-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 13073
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-776-213-20

Query Match 100.0%; Score 723; DB 18; Length 13073;
Best Local Similarity 100.0%; Pred. No. 2, 7e-216;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CTTTCGATTAAGCAAGCAATCACTTGGAGAGAAAAGAGACGCTTGTAAAGGGGAT 120
Db 16 CTTTCGATTAAGCAAGCAATCACTTGGAGAGAAAAGAGACGCTTGTAAAGGGGAT 75
QY 61 ACCATAAAGAGCAAGCAATCACTTGGAGAGAAAAGAGACGCTTGTAAAGGGGAT 120
```

Db 76 ACCATTAAGAGCAGGTAACCTACTTGGAGAAAGAGACCCCTGTGAAGGGGAGAT 135
Qy 121 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGAGCCCGGAGCCCTCT 180
Db 136 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGAGCCCGGAGCCCTCT 195
Qy 181 CTCGCCGACGATTTCTTCTTTCATATCTTCTTTTATCTCTATCCCTTGAAGAAC 240
Db 196 CTCGCCGACGATTTCTTCTTTCATATCTTCTTTTATCTCTATCCCTTGAAGAAC 255
Qy 241 GCATATGACTAATGCTGTGACATCTCCATGCGCTGTGACTTGTGTATCTCAGT 300
Db 256 GCATATGACTAATGCTGTGACATCTCCATGCGCTGTGACTTGTGTATCTCAGT 315
Qy 301 GGTAAAGGACCGTGTGCTGGAAACGTTCTCTGTACAAATCTAAGACAGGGCTACA 360
Db 316 GGTAAAGGACCGTGTGCTGGAAACGTTCTCTGTACAAATCTAAGACAGGGCTACA 375
Qy 361 GTCTCGATAATGAATTAATAGGCAATTTTGTCTAGCGCGCGCGCGCTTTCCCA 420
Db 376 GTCTCGATAATGAATTAATAGGCAATTTTGTCTAGCGCGCGCGCGCTTTCCCA 435
Qy 421 ATAGGAGGCGCGCTTTATGCGGAGCTCTACTTCTTCTATTTGGGTAAAGCCCTTC 480
Db 436 ATAGGAGGCGCGCTTTATGCGGAGCTCTACTTCTTCTATTTGGGTAAAGCCCTTC 495
Qy 481 TGTATTCGCGCAGTGTGCTGTGAGGCTGTGCGCGAGAACTAGTGAATAAGGATGTA 540
Db 496 TGTATTCGCGCAGTGTGCTGTGAGGCTGTGCGCGAGAACTAGTGAATAAGGATGTA 555
Qy 541 TTTTGTGAGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 600
Db 556 TTTTGTGAGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 615
Qy 601 TCTATTAAGAGGAGAAATTTGCTCTACATGTGACAGTTCTGGGACGCTTAACCTT 660
Db 616 TCTATTAAGAGGAGAAATTTGCTCTACATGTGACAGTTCTGGGACGCTTAACCTT 675
Qy 661 TATTGAGAGGACTATCAATCATACAGATATTGTCAAAAAAAGACTAATAATA 720
Db 676 TATTGAGAGGACTATCAATCATACAGATATTGTCAAAAAAAGACTAATAATA 735
Qy 721 CAT 723
Db 736 CAT 738
RESULT 4
US-10-776-213-30
; Sequence 30, Application US/10776213
; Publication No. US20040142478A1
; GENERAL INFORMATION:
; APPLICANT: Astirazeneca AB
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic Acid Expression
; FILE REFERENCE: 3526, 82543
; CURRENT APPLICATION NUMBER: US/10/776, 213
; CURRENT FILING DATE: 2004-02-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 850
; TYPE: DNA
; ORGANISM: *Saccharomyces cerevisiae*
US-10-776-213-30
Query Match 99.6%; Score 720.4; DB 18; Length 850;
Best Local Similarity 99.9%; Pred. No. 4e-216;
Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 CTTGATTTAGCAGCAGACATCATAGACTGCGCTCAATAAATAATACAGGAAA 60
|||||

Db 78 CTTGATTTAGCAGCAGACATCATAGACTGCGCTCAATAAATAATACAGGAAA 137
Qy 61 ACCATTAAGAGCAGGTAACCTACTTGGAGAAAGAGACCGCTGTGAAGGGGAGAT 120
Db 138 ACCATTAAGAGCAGGTAACCTACTTGGAGAAAGAGACCGCTGTGAAGGGGAGAT 197
Qy 121 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGAGCCCGGAGCCCTCT 180
Db 198 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGAGCCCGGAGCCCTCT 257
Qy 181 CTCGCCGACGATTTCTTCTTTCATATCTTCTTTTATCTCTATCCCTTGAAGAAC 240
Db 258 CTCGCCGACGATTTCTTCTTTCATATCTTCTTTTATCTCTATCCCTTGAAGAAC 317
Qy 241 GCATATGACTAATGCTGTGACATCTCCATGCGCTGTGACTTGTGTATCTCAGT 300
Db 318 GCATATGACTAATGCTGTGACATCTCCATGCGCTGTGACTTGTGTATCTCAGT 377
Qy 301 GGTAAAGGACCGTGTGCTGGAAACGTTCTCTGTACAAATCTAAGACAGGGCTACA 360
Db 378 GGTAAAGGACCGTGTGCTGGAAACGTTCTCTGTACAAATCTAAGACAGGGCTACA 437
Qy 361 GTCTCGATAATGAATTAATAGGCAATTTTGTCTAGCGCGCGCGCGCTTTCCCA 420
Db 438 GTCTCGATAATGAATTAATAGGCAATTTTGTCTAGCGCGCGCGCGCTTTCCCA 497
Qy 421 ATAGGAGGCGCGCTTTATGCGGAGCTCTACTTCTTCTATTTGGGTAAAGCCCTTC 480
Db 498 ATAGGAGGCGCGCTTTATGCGGAGCTCTACTTCTTCTATTTGGGTAAAGCCCTTC 557
Qy 481 TGTATTCGCGCAGTGTGCTGTGAGGCTGTGCGCGAGAACTAGTGAATAAGGATGTA 540
Db 558 TGTATTCGCGCAGTGTGCTGTGAGGCTGTGCGCGAGAACTAGTGAATAAGGATGTA 617
Qy 541 TTTTGTGAGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 600
Db 618 TTTTGTGAGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 677
Qy 601 TCTATTAAGAGGAGAAATTTGCTCTACATGTGACAGTTCTGGGACGCTTAACCTT 660
Db 678 TCTATTAAGAGGAGAAATTTGCTCTACATGTGACAGTTCTGGGACGCTTAACCTT 727
Qy 661 TATTGAGAGGACTATCAATCATACAGATATTGTCAAAAAAAGACTAATAATA 720
Db 738 TATTGAGAGGACTATCAATCATACAGATATTGTCAAAAAAAGACTAATAATA 797
Qy 721 CA 722
Db 798 AA 799
RESULT 5
US-10-451-467A-317
; Sequence 317, Application US/10451467A
; Publication No. US20040161840A1
; GENERAL INFORMATION:
; APPLICANT: CONTRERAS, ROLAND HENRI
; APPLICANT: EBERHARDT, INES
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
; APPLICANT: REEMANS, REKA JOSEPHINA
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN YEAST AND FUNGI
; FILE REFERENCE: JAB-1667
; CURRENT APPLICATION NUMBER: US/10/451, 467A
; CURRENT FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: EP 00870318.3
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: EP 01870002.1
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: EP 01870003.9
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 732
; SOFTWARE: PatentIn version 3.1

SEQ ID NO 317
 LENGTH: 680
 TYPE: DNA
 ORGANISM: Saccharomyces cerevisiae
 US-10-451-467A-317

Query Match 68.8%; Score 497.4; DB 18; Length 680;
 Best Local Similarity 99.8%; Pred. No. 7.7e-146;
 Matches 498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 224 ATCCGTTGAAGCAACCGCATATGACTAAATGGTGTGACATCTCCATGGCTGTGACT 283
 Db 1 ATCCGTTGAAGCAACCGCATATGACTAAATGGTGTGACATCTCCATGGCTGTGACT 60
 QY 284 TGTGTATCTCAAGTGTAAAGGCAACCGTGGCTCGGAAACGGTTCCTTGTGACAAAT 343
 Db 61 TGTGTATCTCAAGTGTAAAGGCAACCGTGGCTCGGAAACGGTTCCTTGTGACAAAT 120
 QY 344 CTAGAACAGGGGCTACAGTCTCGATTAATAGAAATAGAGCGCATTTTGTAGCGCGGCC 403
 Db 121 CTAGAACAGGGGCTACAGTCTCGATTAATAGAAATAGAGCGCATTTTGTAGCGCGGCC 180
 QY 404 GCGGCGCCGTTTCCCAATAGGAGCGGCACTTATCGGCGGACCTCTACTTCTCTAT 463
 Db 181 GCGGCGCCGTTTCCCAATAGGAGCGGCACTTATCGGCGGACCTCTACTTCTCTAT 240
 QY 464 TTGGGTAAAGCCCTTCTGTTTGGCGCAAGTGTGTGAGGTGGCGCGGAGAAATA 523
 Db 241 TTGGGTAAAGCCCTTCTGTTTGGCGCAAGTGTGTGAGGTGGCGCGGAGAAATA 300
 QY 524 GTGTAAGGAGATGTAATCTTGATAGAGAAATAGCAAGCGGAAAACTATGCTTAGC 583
 Db 301 GTGTAAGGAGATGTAATCTTGATAGAGAAATAGCAAGCGGAAAACTATGCTTAGC 360
 QY 584 TGGGAGTGTGTTTCAATCATATAAAGGAAATGTTGCTCATACTATGTGACAGTTTC 643
 Db 361 TGGGAGTGTGTTTCAATCATATAAAGGAAATGTTGCTCATACTATGTGACAGTTTC 420
 QY 644 TGGGAGCTCTTAATCTTATTTGAGAGGACATCAATCATACAGATATGTCAAAAAA 703
 Db 421 TGGGAGCTCTTAATCTTATTTGAGAGGACATCAATCATACAGATATGTCAAAAAA 480
 QY 704 AAAAAAGCTAATAATAACA 722
 Db 481 AAAAAAGCTAATAATAAA 499

RESULT 6
 US-10-123-155-198/c
 Sequence 198, Application US/10123155
 Publication No. US20030068794A1
 GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
 APPLICANT: Beresini, Maureen
 APPLICANT: Deforge, Laura
 APPLICANT: Desnoyers, Luc
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Geriltsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Guiney, Austin L.
 APPLICANT: Sherwood, Steven
 APPLICANT: Smith, Victoria
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K
 APPLICANT: Wood, William
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P3330R1C30
 CURRENT APPLICATION NUMBER: US/10/123,155

CURRENT FILING DATE: 2002-04-15
 Prior Application removed - See Palm or File Wrapper
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 198
 LENGTH: 1024
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-123-155-198

Query Match 5.7%; Score 41; DB 14; Length 1024;
 Best Local Similarity 7.7%; Pred. No. 0.088;
 Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

QY 47 TACACTACGAAAAACCATTAAGAGCAAAAGCATCTTGAAGAGAAAAAGGACAG 106
 Db 954 T.CRYAN..MMHY..MMT.HYCD..NYTBD..BH.TH.HB..SN.S.N..YN..NYHS 895
 QY 107 CTGTGAAGGGGAGTGGGGGCTAAGAAAGTCACTTCTTTTCCCTTGGCGGCGGAC 166
 Db 894 GA.T...MM.SBTASD.MKW..WBH.H..SBGAGR.HYBN..YVNS.WHS..A 835
 QY 167 CCGGACCCCTCTCTCCCGACAGATTTCTTCTTCAATCTTCTTATTTCTATC 226
 Db 834 ST..DGDNMC.SGT.K.B.YY..D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
 QY 227 CCGTTAAGCAACCGCATATGACTAATGATGCTGACATCTCCATGSGCTGTACT 286
 Db 774 RN.KYH..ME..R.TYSTTDW..HM..S.RY.....N..RCTYT.S.THH..CTVNS 715
 QY 287 GTGTATCTCAGATGTGTACGACGCGTGGCGGAAAGTTCCTTGTGTGCAATTTCA 346
 Db 714 BS.B.DNY..H..YMY.HRBY.RCA.N.NC..WSCMH.RA.YDD.SVNSBW..T.S 655
 QY 347 GAACAGGGGCTACAGTCTCGATTAATAGAAAGCGATTTTGTGACGCGCGCGG 406
 Db 654 SBDSYNCBB.A...W.RSNN..M.TMAWTS.HR..D.A..YN.TLANC..A.B.RCX 595
 QY 407 GCGCCGTTTCCCAATAGGAGGAGCGGAGTGTGCGGAGCTCATACTTCTTCAATTTG 466
 Db 594 ..AW.HKB...NBRKCNVW.T.S.ANW..HNYTTCRMD.A.RH..CY..BSPCK.NT.KY 535
 QY 467 GGTAAGCCCTTCTGTTTTCGCGCAGTGTGTCAGAGGCTCGCGCGGAGAAACATAGTG 526
 Db 534 ..TTCMRT.WH.BVH.HAT..B..SB..DSST.CH..KB.D.BHA..Y.M.KWNC.R.RH 475
 QY 527 ATAAGGATGTAACCTTTCGATGAGAAATTAGCAAGCGGAAAAAATATGCTGAGCTGG 586
 Db 474 ACB...RDT.C.M.WBH.WMB..AB.HC.W.DG..DB.BKABH..RS.SBSB.H...Y..B 415
 QY 587 GAGTGTGTTTCAATCATATAAAGGAGAAATTTGCTCATACTATGTGACAGTTTCTGG 646
 Db 414 ..S.CWN.KTR.DM.BT..H.M..SR..BM.RH.R...YMA.D.TRHCSCYV..TH.YTR 355
 QY 647 GAGCTTTAATCTTATTTGAGAGGACTATCAAAATCATATGATATTTCAAAAAAAA 706
 Db 354 .YBBSQMAA.YRA.NG.TWT.RRASK..A.S.DHAWCH...BSA.B.KNS.S.RAT. 295
 QY 707 AAGACTAATAATAAC 721
 Db 294 DM.S.M.ATMR...H 280

RESULT 7
 US-10-146-731-198/c

Sequence 198, Application US/10146731
 Publication No. US20030129692A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Beresini, Maureen
 APPLICANT: Deforge, Laura
 APPLICANT: Desnoyers, Luc
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Gao, Wei-Qiang

```

/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P330R1C323
/ CURRENT APPLICATION NUMBER: US/10/146, 731
/ CURRENT FILING DATE: 2002-05-15
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 198
/ LENGTH: 1024
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-146-731-198

Query Match
Best Local Similarity 7.7%; Score 41; DB 15; Length 1024;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

47 TACACTACGGAACCATTAAGAGCAAGCAAGTACTTGTGAGAGAAAGAGACAG 106
954 T.CRYAN..MMHY..MT.HYCD.NYTB.D.BH.TH.H.BB..SN.S.N..YN..NYHS 895
107 CTGTGAAGGGGAGTGGGGCTAAGAGTCACTTCTTCTTCCCTTCCGCGTCCGAC 166
894 GA.T...MM.SBTAASD.MKW..WBH.H..SBCAGR.B.HYBN..YVNSS.WHS..A 835
167 CCGGAGACCCCTCTCTCCCGGACAGATTTCTCTTCAATCTTCTTATTTCTATC 226
834 ST..DGDWNC.SGT.K.B.YY..D.RST.B.SCN.NMW.WC.M.RAATB.M..R.N..N 775
227 CCGTTGAAGCAACGACATGATGATGATGATGATGATGATGATGATGATGATG 286
774 RN.KYH..MH..R.TYSTIDW..HM..S.RY.....N..RCTYT.S.THH..CTVNS 715
287 GTGTATCTCAGAGTGTGAAGGACCGGTGCTCGAAGCGTCTTCTGATCAATTCTA 346
714 BS.B.DNY..H..YVNY.HRNB.Y.RCA.N.NC..WSCNH.RA.YDD.SMNSBW..T.S 655
347 GAACAGGGGCTACAGTCTCATTAATAGATTAATAGAGGCAATTTTGTGACGCGCG 406
654 SBDSYNCB.A...W.RSNM..M.TWAMTS.HR..D..A...YN.TAANC..A.B.RCK 595
407 GCGCCGCTTCCCAATAGAGGAGGCGAGTTATCGGCGAGCTTACTTCTTCTATTTG 466
594 ..AM.HKB...NBRMCNYM.T.S.ANW..HYTTNCRMD.A.RH..CY..BSDCK.NT.KY 535
467 GGTAAAGCCCTTCTGTGTTTTCGCGAGTGTGCTGCGAGCGCGGAGCAATAGTG 526
534 ..TTCCWMT.MH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KMWC.R.RH 475
527 ATAAGGATTAATCTTCATGAGAGATTAAGCAAGCGAGAAATTAATGCTAGCTGG 586
474 ACB...RDT.C.M.WBH.MWB.AB.HC.W.DG..DB.BKABH..RS..SSBS.H...Y..B 415
587 GAGTTGTTTCACTATTAAGAGGAGAAATGTTGCTCAATGCTAGCTATTTCTG 646
414 ..S.CMN.KTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHCSYV.TH.YTR 355
647 GAGCTTAACTTTATGAGAGAGACTCAATCAATCAATGATATTTGCAAAAAAAA 706
354 ..YBBESMAAA.YRA..NG..TWT..RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT.. 295
707 AAGACTAATATAAC 721

```

```

DB 294 DM.S.M.ATWR...H 280

RESULT 8
US-10-140-472-198/C
/ Sequence 198, Application US/10140472
/ Publication No. US2003013888A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: Deforge, Laura
/ APPLICANT: Denoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P330R1C168
/ CURRENT APPLICATION NUMBER: US/10/140, 472
/ CURRENT FILING DATE: 2002-05-06
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 198
/ LENGTH: 1024
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-140-472-198

Query Match
Best Local Similarity 7.7%; Score 41; DB 15; Length 1024;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

47 TACACTACGGAACCATTAAGAGCAAGCAAGTACTTGTGAGAGAAAGAGACAG 106
954 T.CRYAN..MMHY..MT.HYCD.NYTB.D.BH.TH.H.BB..SN.S.N..YN..NYHS 895
107 CTGTGAAGGGGAGTGGGGCTAAGAGTCACTTCTTCTTCCCTTCCGCGTCCGAC 166
894 GA.T...MM.SBTAASD.MKW..WBH.H..SBCAGR.B.HYBN..YVNSS.WHS..A 835
167 CCGGAGACCCCTCTCTCCCGGACAGATTTCTCTTCAATCTTCTTATTTCTATC 226
834 ST..DGDWNC.SGT.K.B.YY..D.RST.B.SCN.NMW.WC.M.RAATB.M..R.N..N 775
227 CCGTTGAAGCAACGACATGATGATGATGATGATGATGATGATGATGATGATG 286
774 RN.KYH..MH..R.TYSTIDW..HM..S.RY.....N..RCTYT.S.THH..CTVNS 715
287 GTGTATCTCAGAGTGTGAAGGACCGGTGCTCGAAGCGTCTTCTGATCAATTCTA 346
714 BS.B.DNY..H..YVNY.HRNB.Y.RCA.N.NC..WSCNH.RA.YDD.SMNSBW..T.S 655
347 GAACAGGGGCTACAGTCTCATTAATAGATTAATAGAGGCAATTTTGTGACGCGCG 406
654 SBDSYNCB.A...W.RSNM..M.TWAMTS.HR..D..A...YN.TAANC..A.B.RCK 595
407 GCGCCGCTTCCCAATAGAGGAGGCGAGTTATCGGCGAGCTTACTTCTTCTATTTG 466
594 ..AM.HKB...NBRMCNYM.T.S.ANW..HYTTNCRMD.A.RH..CY..BSDCK.NT.KY 535
467 GGTAAAGCCCTTCTGTGTTTTCGCGAGTGTGCTGCGAGCGCGGAGCAATAGTG 526

```

Db 534 ..TTCMWRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KNWC.R.RH 475
Qy 527 ATAAAGGATGTAACCTTTCATGAGAGAAATTAAGCAAGCGGAAAAAACTATGCTAGCTGG 586
Db 474 ACB...RDT.C.M.WBH.WWB.AB.HC.W.DG..DB.BKABH..RS.SSBS.H...Y..B 415
Qy 587 GAGTTGTTTTCATCATATAAAGGAGGAAATTTGCTCAGATATGTAAGATTTCTGG 646
Db 414 ..S.CWN.KTTR.DW.BT..H.M..SR.BM.RH.R...YMA.D.TRHGCTY..TH.YTR 355
Qy 647 GAGCTTAACTTTATTTGACAGAGACTATCAATCAATACAGATATTTGTCAAAAAAA 706
Db 354 ..YBBBCSMAA.YRA.NG.TWT..RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
Qy 707 AAGACTAATAATAC 721
Db 294 DW.S.M.ATMR...H 280

RESULT 9

US-10-141-761-198/c
; Sequence 198, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-761-198

Query Match 5.7%; Score 41; DB 15; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.088;

Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

Qy 47 TACACTACGGAAGAAACGATTAAGAGCAAGCAACTTATGGAAGAAAGAGACAG 106
Db 954 T.CRYAN..MMWY..MWT.HYCD.NYTB..BH.TH.H.BB..SN.S.N..YN..NYHS 895
Qy 107 CTGTGAAGGGGAGTGGGGCTAAGATCATTTCTTTCCCTCGCGGCTCGGAGC 166
Db 894 GA.T...WM.SBTAASD.MKW...WBH.H..SBGABR.HYBN...YVYNS.WHS..A 835
Qy 167 CGGGAGCCCTCTCTCCCGGACGATTTCTTCTTCATATATTTCTTTATCTTATC 226
Db 834 ST..DSDMNC.SGT.K.B.YY...D.RST.B..SCN.NNM.WC.M.RAATB.M..R.N..N 775
Qy 227 CCGTGAAGCAACCGGACTAGTAAATGCTGCTGACATCTCCAGTGGGTGAGCTTGT 286
Db 774 RN.KYH..MH..R.TYSTIDW...HM..S.RY.....N...RCTYT.S.THH..CTYNS 715

Qy 287 GTGATATCTCAGATGGTAACGCGACCGGTGGCTCGAAGCGTTCTTCGATATTTCA 346
Db 714 BS.B.DNY...H...YMWY.HRNBV..RCA.N.NC..MSCWH.RA.YDD..SMNSBW..T.S 655
Qy 347 GAACAGGGGCTACAGCTCTGATATATGAAATTAAGCGCATTTTGTAGCGCGCGG 406
Db 654 SBDSYNCB.A...W.RSNR..M.TMAMTS..HR..D..A...YN.TAANC..A.B.RCK 595
Qy 407 GCGCCGTTTCCCATATAGGAGGCGGAGCTTATTCGGCGGAGCTCTACTTCTTCTT 466
Db 594 ..AM.HKB..NBRMCNM.T.S.AMW.HNYTNCMD.A.RH..CY..BSDCK.NT.KY 535
Qy 467 GGTAAAGCCCTTTCTGTTTTCGGCAGGTGCTGCTGACGCTGGCGGAGACATAGTG 526
Db 534 ..TTCMWRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KNWC.R.RH 475
Qy 527 ATAAAGGATGTAACCTTTCATGAGAGAAATTAAGCAAGCGGAAAAAACTATGCTAGCTGG 586
Db 474 ACB...RDT.C.M.WBH.WWB.AB.HC.W.DG..DB.BKABH..RS.SSBS.H...Y..B 415
Qy 587 GAGTTGTTTTCATCATATAAAGGAGGAAATTTGCTCAGATATGTAAGATTTCTGG 646
Db 414 ..S.CWN.KTTR.DW.BT..H.M..SR.BM.RH.R...YMA.D.TRHGCTY..TH.YTR 355
Qy 647 GAGCTTAACTTTATTTGACAGAGACTATCAATCAATACAGATATTTGTCAAAAAAA 706
Db 354 ..YBBBCSMAA.YRA.NG.TWT..RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
Qy 707 AAGACTAATAATAC 721
Db 294 DW.S.M.ATMR...H 280

RESULT 10

US-10-142-885-198/c
; Sequence 198, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C248
; CURRENT APPLICATION NUMBER: US/10/142,885
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-885-198

Query Match 5.7%; Score 41; DB 16; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.088;

Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;


```

QY 47 TACACTACGGAAGAAACCATTAAGAGCAAGCACTACTACTTGGAGGAAAGAGACG 106
DB 954 T.CRYAN..MMMHY..MMT.HYCD..NYTBD..BH.TH.H.BB..SN.S.N..YN..NYHHS 895
QY 107 CTGTGAAGGGGCTAGGAGGCTAAGAGTCACTTCTTCCCTCGCGGCGAC 166
DB 894 GA.T...MM.SBTAASD.MKW..WBH.H..SBCAGR.B.HYEN...YNNSS.MHS..A 835
QY 167 CCGGAGCCCTCTCTCCCGCAGATTTCTTCTTCAATCTTCTTATCTTATC 226
DB 834 ST..DEDMNC.SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CCCTTGAAGAACCGCACTATGACTAATAGTGCTGACATCTTCACTGCTGACTGT 286
DB 774 RN.KYH..MH..R.TYSTIDM..HM..S.RY.....N..RCTYT.S.THH..CTVNS 715
QY 287 GTGTATCTACAGTGTGAAGCAGCAGTGTGCTGGAAGAGCTTCTTCTGACATCTTA 346
DB 714 BS.B.DNY..H..YNNY.HRMBY.RCA.N.NC..WSCNH.RA.YDD.SMNSBW..T.S 655
QY 347 GAACAGGGGCTAGCTCGATATAGATTAATAGAGCAATTTTGTGAGCGCGCG 406
DB 654 SBDSYNCB.A...W.RSNM..M.TWAMTS..HR..D..A...YN.TAANC..A.B.RCK 595
QY 407 GCGCCGCTTCCCAATAGGAGGCGCAGTTTATCGGCGGAGCTTCTTCTTATTTG 466
DB 594 ..AM.HKB...NBRMCNYM.T.S.ANW.HNYTTCRMD.A.RH..CY..BSDCK.NT.KY 535
QY 467 GGTAAAGCCCTTCTGTGTTTGGCAGTGTGCTGAGCGCGCGGAGAACATAGTG 526
DB 534 ..TTCMWRT.WH.BYH.HAT..B..SB.DSST..CH..KB.D.BHA..Y.M.XMWC.R.RH 475
QY 527 ATAAGGATGTAACTTTCGATGAGAGATTAAGAGGCAAGGAAAGAACTATGCTAGCTG 586
DB 474 ACB...RDT.C.M.WBH.WMB..AB.HC.W.DG..DB.BKABH..RS.SBS.H..Y..B 415
QY 587 GAGTTGTTTTCATATATAAAGGAGAAATTTGCTCAGATGTGACATTTCTG 646
DB 414 ..S.CMN.KTTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHGCTY.TH.YTR 355
QY 647 GAGCTTAACTTTTATTCGAGAGACTATCAATCATACAGATATTTGCAAAAAAAA 706
DB 354 .YBBESMAAA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
QY 707 AAGACTAATATAAC 721
DB 294 DM.S.M.ATWR...H 280

RESULT 11
US-10-158-790-198/c
; Sequence 198, Application US/10158790
; Publication No. US2003018079A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerdtisen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME

```

```

; FILE REFERENCE: P3330R1C448
; CURRENT APPLICATION NUMBER: US/10/158,790
; CURRENT FILING DATE: 2002-05-30
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-158-790-198

Query Match 5.7%; Score 41; DB 16; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.088;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

QY 47 TACACTACGGAAGAAACCATTAAGAGCAAGCACTACTACTTGGAGGAAAGAGACG 106
DB 954 T.CRYAN..MMMHY..MMT.HYCD..NYTBD..BH.TH.H.BB..SN.S.N..YN..NYHHS 895
QY 107 CTGTGAAGGGGCTAGGAGGCTAAGAGTCACTTCTTCCCTCGCGGCGAC 166
DB 894 GA.T...MM.SBTAASD.MKW..WBH.H..SBCAGR.B.HYEN...YNNSS.MHS..A 835
QY 167 CCGGAGCCCTCTCTCCCGCAGATTTCTTCTTCAATCTTCTTATCTTATC 226
DB 834 ST..DEDMNC.SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CCCTTGAAGAACCGCACTATGACTAATAGTGCTGACATCTTCACTGCTGACTGT 286
DB 774 RN.KYH..MH..R.TYSTIDM..HM..S.RY.....N..RCTYT.S.THH..CTVNS 715
QY 287 GTGTATCTACAGTGTGAAGCAGCAGTGTGCTGGAAGAGCTTCTTCTGACATCTTA 346
DB 714 BS.B.DNY..H..YNNY.HRMBY.RCA.N.NC..WSCNH.RA.YDD.SMNSBW..T.S 655
QY 347 GAACAGGGGCTAGCTCGATATAGATTAATAGAGCAATTTTGTGAGCGCGCG 406
DB 654 SBDSYNCB.A...W.RSNM..M.TWAMTS..HR..D..A...YN.TAANC..A.B.RCK 595
QY 407 GCGCCGCTTCCCAATAGGAGGCGCAGTTTATCGGCGGAGCTTCTTCTTATTTG 466
DB 594 ..AM.HKB...NBRMCNYM.T.S.ANW.HNYTTCRMD.A.RH..CY..BSDCK.NT.KY 535
QY 467 GGTAAAGCCCTTCTGTGTTTGGCAGTGTGCTGAGCGCGCGGAGAACATAGTG 526
DB 534 ..TTCMWRT.WH.BYH.HAT..B..SB.DSST..CH..KB.D.BHA..Y.M.XMWC.R.RH 475
QY 527 ATAAGGATGTAACTTTCGATGAGAGATTAAGAGGCAAGGAAAGAACTATGCTAGCTG 586
DB 474 ACB...RDT.C.M.WBH.WMB..AB.HC.W.DG..DB.BKABH..RS.SBS.H..Y..B 415
QY 587 GAGTTGTTTTCATATATAAAGGAGAAATTTGCTCAGATGTGACATTTCTG 646
DB 414 ..S.CMN.KTTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHGCTY.TH.YTR 355
QY 647 GAGCTTAACTTTTATTCGAGAGACTATCAATCATACAGATATTTGCAAAAAAAA 706
DB 354 .YBBESMAAA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
QY 707 AAGACTAATATAAC 721
DB 294 DM.S.M.ATWR...H 280

RESULT 12
US-10-137-871-198/c
; Sequence 198, Application US/10137871
; Publication No. US20030207350A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc

```

```
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C153
CURRENT APPLICATION NUMBER: US/10/137, 871
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 198
LENGTH: 1024
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-871-198
```

Query Match 5.7%; Score 41; DB 17; Length 1024;

Best Local Similarity 7.7%; Pred. No. 0.086; Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

```
47 TACACTACGAAAAACCATTAAGAGCAAGATCTACTTGGAGAAAGAGACG 106
954 T.CRYAN..MMMHY..MWT.HYCD.NYTB.D..BH.TH.H.BB..SN.S.N..YN..NYHS 895
107 CTGTGAAGGGGAGTGGGGCTAAGAGTCACTTCTTCTCCGCGCGGAC 166
894 GA.T...MM.SBTAASD..MKM..H..SBQAGR.B..HYBN...YNYNS.WHS..A 835
167 CCGGAGACCCCTCTCTCCCGCAGATTTCTCTTCATATCTTCTTATTCATC 226
834 ST..DGMNC.SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
227 CCGTTGAAGACCGCATATGACTAATGCTGACATCTTCATGCTGACTGT 286
774 RN.KYH..MH..R.TYSTTDW...HM..S.RY.....N..RCYT.S.THH..CTYNS 715
287 GTGTATCTACAGTGGTAAGCGGACGCTGCTGGAAACGTTCTTCTGACATTTCA 346
714 BS.B.DNY...H..YNNY.HRNB.Y.RCA.N.NC..WSCMH.RA.YDD..SMNSBW..T.S 655
347 GAACAGGGGCTACAGTCTCGATATAGATAATTAAGCGCATTTTGTGACGCGCGCG 406
654 SBDSYNCB.A...W.RSNN..M.TMAWTS.HR..D..A...YN.TAANC..A.B.RCK 595
407 GCGCCGCTTCCCAATAGGAGCGGAGTTATGCGGAGCTCTACTTCTTCTATTG 466
594 ..AM.HKB...NBRMCNYM.T.S.ANW.HNYTNCRMD.A.RH..CY..BSDCK.NT..KY 535
467 GGTAAAGCCCTTCTGTGTTGCGCAGTGTGCTGACAGCTGCGCGGAGAACATAGTG 526
534 ..TTCMRT..MH..BYH.HAT..B..SB.DSST..CH..KB..D.BHA..Y.M.KXWC.R..RH 475
527 ATAAGGATGTAACTTTCATGATGAGATTAAGCAAGCGGAAAAAATGATGCTAGCTGG 586
474 ACG...RDT.C.M.WBH.WMB.AB.HC.W.DG..DB.BRABH..RS.SBS.H..Y..B 415
587 GAGTTTCTTCAATATATAAAGGAGAAATGTTGCTCACTATGTAAGCTTTCTGG 646
414 ..S.CWN..KTRR.DM..BT..H.M..SR..BM.RH.R...YMA.D.TRHSCYV..TH.YTR 355
647 GACGTCTAATCTTTATGTAAGAGGAGCTATCAATGATCAAGATATGTCACAAAAA 706
354 .YBBSCMAAA.YRA.NG.TWT.RRASK..A.S.DHABWCH..BSA.B.KUS.S.RAT.. 295
```

```
QY 707 AGACTAATAATAAC 721
DB 294 DM.S.M.ATWR...H 280
```

RESULT 13

US-10-140-923-198/c
Sequence 198, Application US/10140923
Publication No. US20030207355A1
GENERAL INFORMATION:

```
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C18
CURRENT APPLICATION NUMBER: US/10/140, 923
CURRENT FILING DATE: 2002-05-07
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 198
LENGTH: 1024
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-923-198
```

Query Match 5.7%; Score 41; DB 17; Length 1024;

Best Local Similarity 7.7%; Pred. No. 0.086; Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

```
47 TACACTACGAAAAACCATTAAGAGCAAGATCTACTTGGAGAAAGAGACG 106
954 T.CRYAN..MMMHY..MWT.HYCD.NYTB.D..BH.TH.H.BB..SN.S.N..YN..NYHS 895
107 CTGTGAAGGGGAGTGGGGCTAAGAGTCACTTCTTCTCCGCGCGGAC 166
894 GA.T...MM.SBTAASD..MKM..H..SBQAGR.B..HYBN...YNYNS.WHS..A 835
167 CCGGAGACCCCTCTCTCCCGCAGATTTCTTTCATATCTTCTTATTCATC 226
834 ST..DGMNC.SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
227 CCGTTGAAGACCGCATATGACTAATGCTGACATCTTCATGCTGACTGT 286
774 RN.KYH..MH..R.TYSTTDW...HM..S.RY.....N..RCYT.S.THH..CTYNS 715
287 GTGTATCTACAGTGGTAAGCGGACGCTGCTGGAAACGTTCTTGTGACATTTCA 346
714 BS.B.DNY...H..YNNY.HRNB.Y.RCA.N.NC..WSCMH.RA.YDD..SMNSBW..T.S 655
347 GAACAGGGGCTACAGTCTCGATATAGATAATTAAGCGCATTTTGTGACGCGCGCG 406
654 SBDSYNCB.A...W.RSNN..M.TMAWTS.HR..D..A...YN.TAANC..A.B.RCK 595
407 GCGCCGCTTCCCAATAGGAGCGGAGTTATGCGGAGCTCTACTTCTTCTATTG 466
594 ..AM.HKB...NBRMCNYM.T.S.ANW.HNYTNCRMD.A.RH..CY..BSDCK.NT..KY 535
```

```

QY 467 GGTAAAGCCCTTCTGTTTCGCGCAGTGGTGTGAGGCTGCGCGGAGAAATAGTG 526
DB 534 ..TTGMMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KWMC.R.RH 475
QY 527 ATAGAGGATTAACCTTTCATGAGAGATTAAGCAGCGGAAAAAACTAGCTAGCTGG 586
DB 474 ACB...RDT.C.M.WBH.WWB.AB.HC.W.DG..DB.BKABH..RS.SSBS.H...Y..B 415
QY 587 GAGTGTGTTTCAATATATAAAGGAGAAATGTTGCTCACTATGTGACAGTTTCTGG 646
DB 414 ..S.CWN.KTTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHGSCY.YH.YTR 355
QY 647 GAGGTCTTAACCTTTATGAGAGAGACTATCAATCATAGATATTTGTAACAAAAA 706
DB 354 .YBBSCMAAA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
QY 707 AAGACTAATATAAC 721
DB 294 DM.S.M.ATMR...H 280

```

```

RESULT 14
US-10-141-756-198/c
; Sequence 198, Application US/10141756
; Publication No. US20030207359A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C200
; CURRENT APPLICATION NUMBER: US/10/141,756
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-756-198

```

```

Query Match 5.7%; Score 41; DB 17; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.088;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;
QY 47 TACACTACGAAAGAAACATTAAGAGCAAGGATACCTACTTGGAGGAAAGAGACAG 106
DB 954 T.CRYAN..MMWHY..MWT.HYCD..NYTBD..BH..TH.H.BB..SN.S.N..YN..NYHS 895
QY 107 CTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 166
DB 894 GA.T...MM.SPTASD.MKW...WBH.H..SBGAGR.BHYN...YANNS.WHS..A 835
QY 167 CCGGAGCCCTCTCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 226
DB 834 ST..DGDWNC.SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CGCTTAAGCAACGCACTATGACTAATAGTGTGAGCAATCTCCATGCTGTGACTGTG 286

```

```

DB 774 RN.KTH..MH..R.TYSTDM...HM..S.RY.....N...RCYTT.S.THH..CTYNS 715
QY 287 GTGTATCTCAGATGTGTAACGCGACCGTGTGCGAAGCGTTCTTCTGTGCAATTTCTA 346
DB 714 BS.B.DNY...H...YMWY.HNBY.RCA.N.NC..WSCMH..RA.YDD.SPANSBW..T.S 655
QY 347 GAACAGGAGGCTACAGTCTGATTAATGAAATTAATAGCGGATTTTCTAGCGCGCGG 406
DB 654 SBDSYNCB.A...W.RSNN..M.TMAMTS..HR..D..A...YN.TAANC..A.B.RCK 595
QY 407 GCGCCGTTTCCCAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 466
DB 594 ..AW.HKB...NBRKCNM.T.S.ANW..HNYTNCMD.A.RH..CY..BSPDK.NT.KY 535
QY 467 GGTAAAGCCCTTCTGTTTCGCGCAGTGGTGTGAGGCTGCGCGGAGAAATAGTG 526
DB 534 ..TTGMMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KWMC.R.RH 475
QY 527 ATAGAGGATTAACCTTTCATGAGAGATTAAGCAGCGGAAAAAACTAGCTAGCTGG 586
DB 474 ACB...RDT.C.M.WBH.WWB.AB.HC.W.DG..DB.BKABH..RS.SSBS.H...Y..B 415
QY 587 GAGTGTGTTTCAATATATAAAGGAGAAATGTTGCTCACTATGTGACAGTTTCTGG 646
DB 414 ..S.CWN.KTTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHGSCY.YH.YTR 355
QY 647 GAGGTCTTAACCTTTATGAGAGAGACTATCAATCATAGATATTTGTAACAAAAA 706
DB 354 .YBBSCMAAA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
QY 707 AAGACTAATATAAC 721
DB 294 DM.S.M.ATMR...H 280

```

```

RESULT 15
US-10-141-759-198/c
; Sequence 198, Application US/10141759
; Publication No. US20030207361A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C197
; CURRENT APPLICATION NUMBER: US/10/141,759
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-759-198

```

```

Query Match 5.7%; Score 41; DB 17; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.088;

```

```

Db      294 DM.S.M.ATMR...H 280

```

Job time : 529 secs